

Qy	61	SSTWSTFTTFFQISTPYTSPGGCLAFFIAPYDTVTPPSAGNGLGFLPNLNA LRNS	120
		: : : : : : : : : : : : : : : : : : : :	
Dd	61	DSAVLSFDRIINFEISTPYSRIA DGLAFFIAPPDSVI--SVHGFGCLFPNPANTLN NS	118
		: : : : : : : : : : : : : : : : : : : :	
Qy	121	TTSKETTTIDVNAASNVAVEEDTYPNDIGNDGYRKHIGCIDVNSIRSKATVAMDQNCKYT	180
		: : : : : : : : :	
Dd	119	STSENQT-TTKAASNVAVEEDTYLPNDYGDPNYTHIGIDVNSIRSKVTAKWDNQNKCI	177
		: : : : : : : : : : : : : : : : : : : :	
Qy	181	ATAHISYNSAKRLSNTVPYPGCGKAVSLSHDVLTOLVQLPWLRVFESASTGLEK	234
		: : : : : : : : : : : : : : :	
Dd	178	ATAHISYNSVKRLSYTYAGSKPATLSVDIELHTVLPEWRVGLSASTGDQK	231
		: : : : : : :	

RESULT 2
 US-08-881-189B-23
 ; Sequence 23, Application US/08881189B
 ; Patent No. 6310195
 ; GENERAL INFORMATION:
 ; APPLICANT: Colucci et al.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
 ; PROGENITOR CELL PRESERVATION FACTOR
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11753
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/881,189B
 ; FILING DATE: June 24, 1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 286 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 US-08-881-189B-23

db 200 ATAHISYNSVKRLSVTSYYAGSKPATLSYDIELHTVLPENVRVGLSASTGQDK 253

```

RESULT 3
US-09-141-821-2
; Sequence 2, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSTAI
; APPLICANT: Szuzsanna Magdalena BARDOCZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTÉLES
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-2

Query Match          37.8%; Score 457; DB 3: Length 285;
Best Local Similarity 44.7%; Pred. No. 2e-41;
Matches 105; Conservative 40; Mismatches 68; Indels 22; Gaps

QY      3 SVSTFTTKFDSDKDLFMFOGHT-ISSSNVTLQLTKLDSNGPNPVSIVGRVLYSAPLRWES 61
DB      34 SLSPFPFKHSQDPLIFOSDALVTSGVLQLTV--NDRGVYDSIGRVLAAAPFIQWDS 91
        :|||:||| |||:||| ::|||::| | |||:|||:|||::|
QY      62 ST-VYSTETTTTQTISTPYTPSPGDGLAFLLAPYDIVIPNSAGNLLGLFPNLALRNS 120
DB      :|||:|||:||| ||| ||| ||| ||| ||| ||| ||| |||
        :|||:|||:||| ||| ||| ||| ||| ||| ||| ||| |||
QY      121 TTSKETTIDVNNAASNNVVAVEDFYPDNIIGDPYRKHGTIDVNSIRSKATVAMDWONGKT 180
DB      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
        :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      145 SYNK-----SNQIVAVEFDITFRN-VAMDPNGIIMHGIDVNSIQSVRVRDWANGEV 194
DB      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
        :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      181 ATAHISYNASAKRLSVTFYPFG-KGAVSLSHDVELTOVLQPIWRVGFSASTGLEK 234
DB      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
        :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      195 ANVETSYFAQPKTSASTAYVESLSPEKSFITSATVDIKKVIPEWVRVGTTATTGLSE 249
DB      :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
US-09-141-821-1
; Sequence 1, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szuzanna Magdolna BARDOCCZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELER
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows version 3.0
; SEQ ID NO 1
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-1

```


[illegible]

Db 31 EGISFNFTNFTRGQGVLLGOANIMANGILALT---NHTNP-TWNTGRALYSKVPVIMD 86
QY 61 SST-VVSTFTFTFOISTPTSPGDLAFPLAPYDVIPNSAGNLLGLFPNLAURN 119
Db 87 SATGNVASFVTSFVKEIGGIPADGIVFPLAP-EARIPDNSAGGOLGI----- 136
QY 120 STTSKETTIDVNAASNNVAVFDTYPNDNIGDPYRKHGIDVNSIRSKATVADWQNGK 179
Db 137 -----VNANKAYNPFVGFEDTYSNN--WDPKSAHGIDASSLSLRTVKWKNVSGS 186
QY 180 TATAHISYNSASKRLSVTTFPGGKAVSLSDHVELTQVLPQWIRGVFSAST 230
Db 187 LVKVSIIYLSLKTLSVVVTHENGQISTIAQVLDKAVLGEKVRVGFTAAT 237

RESULT 10

US-08-881-189B-13
; Sequence 13, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hofmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-13

Query Match 23.6%; Score 285.5; DB 4; Length 105;
Best Local Similarity 58.4%; Pred. No. 1.7e-23;
Matches 66; Conservative 13; Mismatches 21; Indels 13; Gaps 4;

QY 1 AQSVSFTFTKDSQKDLMPFGHTISSNNVQLTKLDSNGNPVSTSVGRVLYSAPRLWE 60
Db 1 AQSLSFSTKFDPNQEDLIQGTATS-----KLSAGNPVSSSAGRVLYSAPRLWE 52
QY 61 SSTVSTFTFTFOISTPTSPGDLAFPLAPYDVIPNSAGNLLGLFPN 113
Db 53 DSAVLTSFDP--TIYIFNTSRIADGLA-FIAPPDVSI--SYHGGLGLFPN 100

RESULT 11

US-09-228-986-77

; Sequence 77, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated from plant cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 77
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-77

Query Match 21.0%; Score 254; DB 4; Length 632;
Best Local Similarity 31.8%; Pred. No. 6.9e-19;
Matches 81; Conservative 37; Mismatches 91; Indels 46; Gaps 10;

QY 1 AOSV-----SFTFTKDSQKDLMPFGHTISSNNVQLTKLDSNGNPVSTSVGRVLYS 53
Db 23 AQAVEDRRHDTTFLDFGNGTNLILEANASVIGSESVLSLT-----NHSHEFMLGRALYA 77
QY 54 APLRLWESSIVVSTFTFTFOISTPTSPGDLAFPLAPYDVIPNSAG--NLGLLF 111
Db 78 APVQMKNNHT-VSSFTFTFVSIVPPSPNEGGLAFIMTPYTS---PMGAQPVQYLG- 132
QY 112 PNLNALRNSSTSKETTIDVNAASNNVAVFDTYPNDNIGDPYRKHGIDVNSIRSKAT- 170
Db 133 --LNLTSN-----QOPYNHLEFAVEFTIMNVEFKDPDRNHVGVDSINLSISVOTE 179
QY 171 VAWDW-----QNGKTATAHISYNSASKRLSVT---TFYPGGKAVSLSDHVELTQV 217
Db 180 TAGYNGGEEFHELNLRSGRNIAQWIDYDHESSLNVITVAGLPRPQRLISLQIDLQNI 239
QY 218 LPQWIRVGFSAATGL 232
Db 240 VEEKMLVGFSAATGL 254

RESULT 12

US-08-853-659A-43
; Sequence 43, Application US/08853659A
; Patent No. 5925522
; GENERAL INFORMATION:
; APPLICANT: Wong, K.K.; Saffer, J.D.
; TITLE OF INVENTION: A Salmonella Sequence, Methods Of Detection
; TITLE OF INVENTION: Of A
; TITLE OF INVENTION: Salmonella Sequence, And Methods Of Detection Of Salmonella
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Paul W. Zimmerman
; ADDRESSEE: Intellectual Property Services
; ADDRESSEE: Battelle Memorial Institute
; ADDRESSEE: PNNL P.O. Box 999
; STREET: Washington Way
; CITY: Richland
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 99352
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.50", 1.44 Mb storage
; COMPUTER: IBM PC/XT/AT
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Processor (WordPerfect 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,659A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: none
; FILING DATE: n/a
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 252 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-853-659A-43

Query Match
Best Local Similarity 8.0%; Score 96.5; DB 2; Length 252;
Matches 52; Conservative 26; Mismatches 72; Indels 69; Gaps 12;

QY 17 DLMOGHFTISSN--VIOLTKLDNGNPGVSTSGVRLYSAPLRLWSSVWSTFTET---71
Db 25 DTGQGDNNHTSTQTFALQHIDDDAARVTVSV-----EHGGVTTTFDQTKGT 72

QY 72 --FTFQISTPYTS-PPGDLGAFFLAPYD-TVIPPSAGNLLGLFPNLRNLTSTSKETT 127
Db 73 GGWTF---TPPTSWADGD-----YTLVSVEDKAGN-----TSHSASLTVT 110

QY 128 IDVNAASNNVAVEDTYPNDNIGDPYRKHIGI-----DVSIR-----SKA 169
Db 111 VDTQTAINNIELVDSGIPDDNLTNNRPHFQVTVPTDVNVVRLSIDGKTNFNATQSAT 170

QY 170 TVAMDW-----QNGKTATAHISYNSASKRLSVTTFYP 201
Db 171 PGVDYIWPDDVADGGYTLTVEAT-DEAGNKATQTSISP 208

RESULT 13
US-09-134-001C-4463
; Sequence 4463, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4463
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4463

Query Match
Best Local Similarity 7.9%; Score 95; DB 4; Length 2137;
Matches 58; Conservative 44; Mismatches 111; Indels 40; Gaps 12;

QY 1 AQSVSFTEFKDSQKDLFMFGHTISS--NVIQLTKLDNGNPGVSTSGVRLYSAPLRL- 58
Db 610 SQTVIYKFKDVGQPOISVDSQTSQREVSKTINPTITTTDNSKDLVLTIV---TGLPSGLS 665

QY 59 --WESSTVSTFTETFTFOISTPYTSPPGD--GLAFLAPYDVTIP-----PNSAGNLLG 109
Db 666 FDOTNTITGTPSEVGTITVTNTTATGNVTISKQPTITIQDTISFVNVTPOASE---722

QY 110 LPPNLRNLTSTKETTIDVNAASNNVAVEDTYPNDNIGDPYRKHIGIDVNSIRSKA 169
Db 723 VTFPNTPIITADNSGKVTVHTVGLPGLKFDASTNSIVGTPTQ-----IGTNTITIES 778

QY 170 TVAMDQNGKTATAHISY-----NSASKRL-----SVTTFPGCKAVSLSHDVELAQVL- 218
Db 779 TDA----SGNKTTKINIEVTRNSASDSTSIIVNSVSTSI--SNSTSLSDSVKASQSL 832

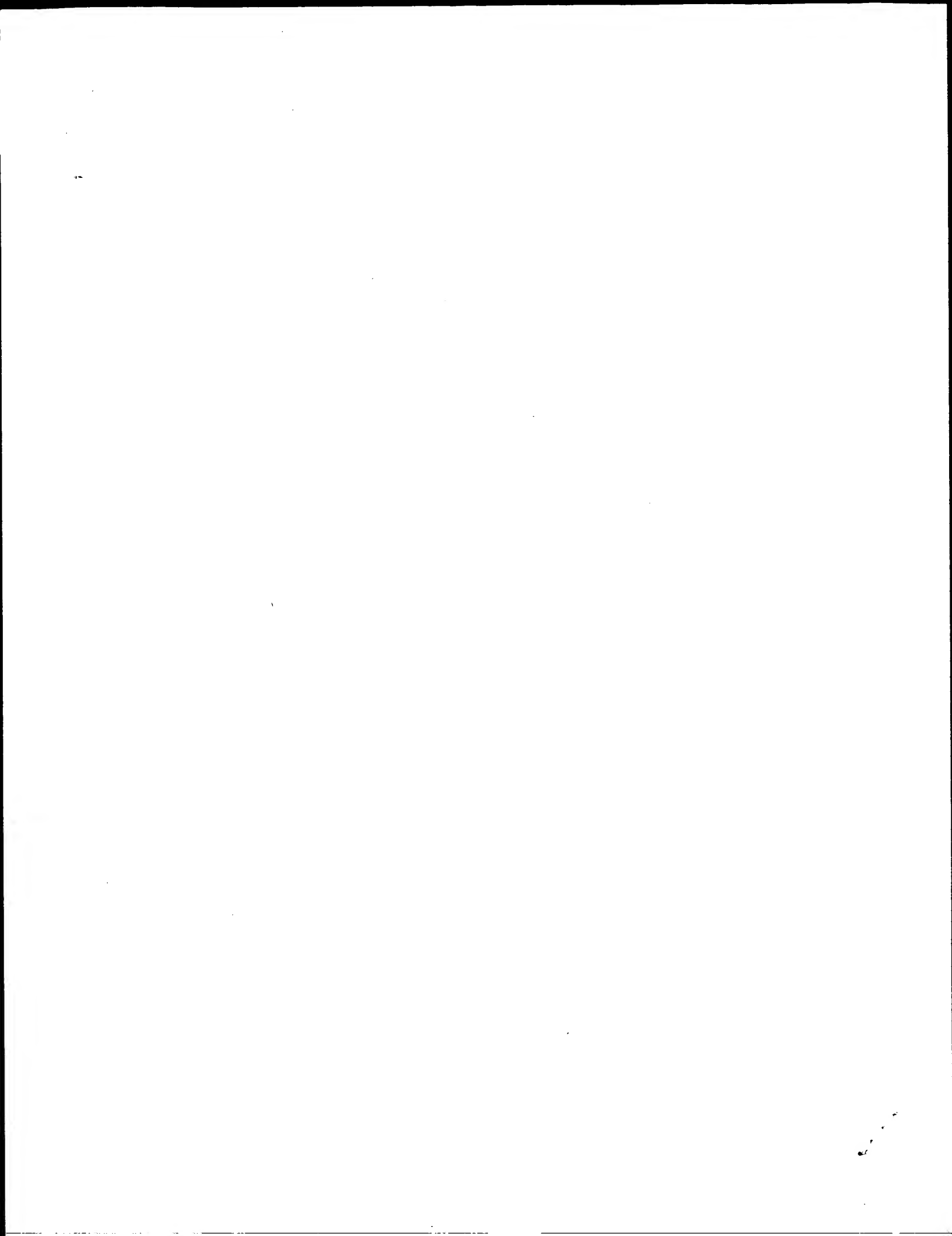
QY 219 -PQWIRVGFSA 230
```

```
Db 833 TSKSLSESLSAST 845
; : : ||||
RESULT 14
US-08-856-253-7
; Sequence 7, Application US/08856253
; Patent No. 6288214
; GENERAL INFORMATION:
; APPLICANT: Hook, Magnus
; APPLICANT: Patti, Joseph M.
; APPLICANT: House-Pompeo, Karen
; APPLICANT: Sthanam, Narayana
; APPLICANT: Symersky, Jindrich
; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: U.S.
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,253
; FILING DATE: Concurrently Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,678
; FILING DATE: 16-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TAMK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-856-253-7

Query Match
Best Local Similarity 7.8%; Score 94.5; DB 4; Length 345;
Matches 59; Conservative 32; Mismatches 73; Indels 81; Gaps 16;

QY 3 SVSTFTFKDSQKDLFMFGHTISS--NVIQLTKLDNGN--PVSTSGVRL-----51
Db 107 NVIYFTDYVNTKDDVK---ATLTPAYIDPENVKKTGNVTLATIGSTANKTVLVDYE 163
QY 52 -YSAPLRLWESSTVSTFTETFTFOISTPYTSPPGDGLAFLAPYDVTIPPSAGNLLGL 110
Db 164 KYGFYNLSIKGTIDQIDKTNNTYR-QTIYVNPSTG-----DNVIAPVLTGNLK--210
QY 111 FPNL--NAL--RNSSTSKETTIDVNA-----SNNVAVE 141
Db 211 -PNTDSNALIDQNTSIKVKVD-NAADLSBSYFVNPFENFEDVTNSVNTITFPNQYKVE 268
QY 142 PDTVPNDNIGDPYRKHIG--IDVNS-----KATVAV---DWQNGKTATAHIS 186
Db 269 FNT-PDQITTPYIVVNGHIDPNKSGDLAURSTLYGYNSTIWRSMWDN-----EVA 321
QY 187 YNSAS 191
; : : :
```

Search completed: February 26, 2003, 16:41:27
Job time : 13.9775 secs



Db 36 SDSLSEFNFPNSENDELIFQKASISSNETLETRISSQOPATSSVGRALYTPVRLW 95
 QY 60 ESST-VVSTFETFTFOISTPTSPGDDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 118
 Db 96 DKSTGLASPKTTFSAITSP-TQDGDGGAFFITAPDPT--PGYGGGLLGLFNGFN-LR 151
 QY 119 NSTTSKETTIDVNAASNNVAVVEFDYTPNDNIGDPYRKHHIGIDVNSIRSKATVAMDQNG 178
 Db 152 NSSNN--GVAVNNQSAQIVAVEFDYINGQC-DPKYRHVGIDVNSITSLAYTQWQNG 207
 QY 179 KTATAHISYNSAKRLSVTTFTYPGGKAVSLSHDVELTQVLPOWIRVGFSASTG 231
 Db 208 VKATAQISYNPASOKLTAVTSYNPSTPLTVSLDIDLQTLVLPWVRVGFSASTG 260

RESULT 5

JC2268
 hemagglutinin - Maackia amurensis
 C:Species: Maackia amurensis
 C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 07-May-1999
 C:Accession: JC2268
 R:Konami, Y.; Ishida, C.; Yanamoto, K.; Osawa, T.; Irimura, T.
 J. Biochem. 115, 767-777, 1994
 A:Title: A unique amino acid sequence involved in the putative carbohydrate-binding domain of hemagglutinin (MAH).
 A:Reference number: JC2268; MUID:94375425; PMID:8089095

A:Accession: JC2268
 A:Molecule type: protein
 A:Residues: 1-247 <KON>
 A:Experimental source: seed
 C:Superfamily: plant lectin
 C:Keywords: calcium binding; glycoprotein; manganese
 F:111,177,189/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:125,127,138,143/Binding site: manganese (Glu, Asp, His) #status predicted
 F:127,135,138/Binding site: calcium (Asp) #status predicted

Query Match 39.7%; Score 480.5; DB 2; Length 247;
 Best Local Similarity 44.8%; Pred. No. 1.2e-32;
 Matches 107; Conservative 38; Mismatches 71; Indels 23; Gaps 7;

QY 1 AOSVSFTTKFSDQKDLMPQG-HTSSNVIOQLTKLDSNGNPVSTSVGRVLYSAPLRLW 59
 Db 1 SDELSFTINNFVPEADLLFQGEASVSSTGLQTRVE-NGQPQKYSVGRALYAAPVRW 59
 QY 60 ESST-VVSTFETFTFOISTPTSPGDDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 118
 Db 60 DSTTGAVASFSTFTFVKA--ARGASDGLAFFLAPDSDQIPSGSVSKYLGLFNNSN--- 114
 QY 119 NSTTSKETTIDVNAASNNVAVVEFDYTPNDNIG--DPYRKHHIGIDVNSIRSKATVAMDQ 176
 Db 115 -----SDSSNQIVAVEFDYFGHSYDPDPNRYRHIGIDVNGIESIKTVQWDWI 162
 QY 177 NGKTATAHISYNSAKRLSVTTFTYPGGK-AVLSHDVELTQVLPOWIRVGFSASTGLEK 234
 Db 163 NGGVAFAFATITLAPNKTILASLVYPSNQITFSAASVDLKEILPEWVRVGFSAATGAPK 221

RESULT 6

JC5444
 leukaagglutinin precursor - Maackia amurensis
 C:Species: Maackia amurensis
 C:Date: 17-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 07-May-1999
 C:Accession: JC5444; PC4324
 R:Yamamoto, K.; Konami, Y.; Irimura, T.
 J. Biochem. 121, 756-761, 1997
 A:Title: Sialic acid-binding motif of Maackia amurensis lectins.
 A:Reference number: JC5444; MUID:97306060; PMID:9163528

A:Accession: JC5444
 A:Molecule type: mRNA
 A:Residues: 1-287 <YAM>
 A:Accession: PC4324
 A:Molecule type: protein

A:Residues: 30-55;74-176;187-253 <YA2>
 A:Experimental source: seed
 C:Comment: This protein is a leguminous lectin. It interacts with high affinity with te galactosyl residues.
 C:Superfamily: plant lectin
 C:Keywords: glycoprotein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-287/Product: leukaagglutinin #status predicted <MAT>
 F:90,142,208,220/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:272/Disulfide bonds: interchain #status predicted

Query Match 39.7%; Score 479.5; DB 2; Length 287;
 Best Local Similarity 43.6%; Pred. No. 1.8e-32;
 Matches 103; Conservative 43; Mismatches 69; Indels 21; Gaps 6;

QY 1 AOSVSFTTKFSDQKDLMPQG-HTSSNVIOQLTKLDSNGNPVSTSVGRVLYSAPLRLW 59
 Db 30 SDELSFTINNFVPEADLLFQGEASVSSTGLQTRVE-NGQPQKYSVGRALYAAPVRW 88
 QY 60 ESST-VVSTFETFTFOISTPTSPGDDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 118
 Db 89 DNTTGSVASFSFTFVVKAPNPDIITSDGLAFYLPDSDQIPSGSVSKYLGLFNNSN--- 145
 QY 119 NSTTSKETTIDVNAASNNVAVVEFDYTPNDNIGDPYRKHHIGIDVNSIRSKATVAMDQ 176
 Db 146 -----SDSSNQIVAVEFDYFAHSYDPDPNRYRHIGIDVNGIESIKTVQWDWI 193
 QY 177 NGKTATAHISYNSAKRLSVTTFTYPGGKAV-SLSDHVELTQVLPOWIRVGFSASTG 231
 Db 194 NGGVAFAFATITLAPNKTILASLVYPSNQITFSAASVDLKEILPEWVRVGFSAATG 249

RESULT 7

S66355
 lectin-related storage protein precursor - Cladrastis lutea (fragment)
 C:Species: Cladrastis lutea
 C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 20-Aug-1999
 C:Accession: S66355; S66299
 R:van Damme, E.J.W.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
 Plant Mol. Biol. 29, 579-598, 1995
 A:Title: A lectin and a lectin-related protein are the two most prominent proteins in A:Reference number: S66299; MUID:96123235; PMID:8534854

A:Accession: S66355
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-290 <VAN>
 A:Cross-references: EMBL:U21940; NID:g1141754; PID:AAC49150.1; PID:g1141755
 A:Accession: S66299
 A:Molecule type: protein
 A:Residues: 37-56 <VAM>
 A:Experimental source: bark
 C:Comment: This lectin-related protein has no carbohydrate binding activity.
 C:Superfamily: plant lectin
 C:Keywords: lectin

F:1-36/Domain: signal sequence (fragment) #status predicted <SIG>
 F:37-290/Product: lectin-related storage protein #status experimental <MAT>

Query Match 39.5%; Score 477.5; DB 2; Length 290;
 Best Local Similarity 45.1%; Pred. No. 2.7e-32;
 Matches 105; Conservative 42; Mismatches 63; Indels 23; Gaps 8;

QY 2 QSVSFTTKFSDQKDLMPQGHT-ISSNVIOQLTKLDSNGNPVSTSVGRVLYSAPLRLWE 60
 Db 38 EALSFTTKFVSNQDELLQDALVSSKGEQLQTRVE-NGQPPIHVSVGRALYSDPVHWDI 96
 QY 61 SST-VVSTFETFTFOISTPTSPGDDGLAFFLAPYDVTIPPNSAGNLLGLFPNLNLR 119
 Db 97 SSTGSVASEVTSFTFVVEAPNENKTADGLAFFLAPDITQV--QSLGGFGLGF-----N 147
 QY 120 STTSKETTIDVNAASNNVAVVEFDYTPNDNIGDPYRKHHIGIDVNSIRSKATVAMDQNGK 179
 Db 148 SS-----VYNSSNQILAVEFDFTSNS--WDPTARHIGIDVNSIESTRTATGWGRNCE 197

QY 232 LE 233

Db 245 DE 246

RESULT 11

A25701

phytohemagglutinin chain L precursor - kidney bean

N:Alternate names: PHA-L

C:Species: Phaseolus vulgaris (kidney bean)

C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 11-Apr-1995

C:Accession: A25701

R:Voelker, T.A.; Staswick, P.; Chrispeels, M.J.

EMBO J. 5, 3075-3082, 1986

A:Title: Molecular analysis of two phytohemagglutinin genes and their expression in Phaseolus vulgaris

A:Reference number: A25701

A:Accession: A25701

A:Molecule type: DNA

A:Residues: 1-273 <VOE>

A:Experimental source: cv. Pinto UI111

C:Genetics:

A:Gene: dlec2

C:Superfamily: plant lectin

Query Match

Best Local Similarity 37.0%; Score 447; DB 2; Length 273;

Matches 101; Conservative 41; Mismatches 72; Indels 24; Gaps 7;

QY 1 AQSVSFTFKFDS-DOKDLMFQGH-TISSNVQLTKLDSNGNPVSTSGRVLVSAPLRL 58

Db 19 ANSASQTFSEDFRNETLILQDGSVSSGQLRLTNVNSNGEPTVGLGRAFYSAPTOI 78

QY 59 WESST-VSTFTFTFTFOISTPTSPGDLAPFLAPYDTVIPNSAGNLLGLFPNLNAL 117

Db 79 WDTYTCNVASEFDNFENILVPNAGPADGLAFALVPVGS--QPKDKGGFLGLFDGSGN-- 134

QY 118 RNSTTKETIDVNAASNNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWQN 177

Db 135 -----SNFHTVAVEFDLYNKD-WDPREHIGIDVNSIKSIKTPWDFVN 178

QY 178 GKTATAHISYNSAKRLSVTTFPGKG-AVLSHDVELTOVLQWIRVGFSASTGLEK 234

Db 179 GENAEVHITYESTKLLVASLVYPSLKTSTVSDTVDLKSVLPFWVSGFSATTGITK 236

RESULT 12

S62690

agglutinin I precursor - black locust

N:Alternate names: lectin

C:Species: Robinia pseudoacacia (black locust)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S62690; S62685

R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.

Plant Mol. Biol. 29, 1197-1210, 1995

A:Title: The seed lectins of black locust (Robinia pseudoacacia) are encoded by two genes

A:Reference number: S62685; MUID:96191285; PMID:8616218

A:Accession: S62690

A:Molecule type: protein

A:Status: nucleic acid sequence not shown

A:Residues: 1-285 <VAV>

A:Cross-references: EMBL:U24249

A:Experimental source: seed

A:Accession: S62685

A:Molecule type: protein

A:Residues: 32-51 <VAV>

C:Superfamily: plant lectin

C:Keywords: glycoprotein; homotetramer; lectin

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-285/Product: agglutinin I #status experimental <MAT>

Query Match

Best Local Similarity 36.8%; Score 445; DB 2; Length 285;

Matches 100; Conservative 41.9%; Mismatches 56; Indels 22; Gaps 8;

Matches

98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

QY 3 SVSFTTTFKDSQDKLMFQGH-TISSNVQLTKLDSNGNPVSTSGRVLVSAPLRLWES 61

Db 34 SLSEFPKFAFPNQPYLIFQDALVTSTGVLQLTNV-VNGVPPRRSIGRALYAAFFQIWDN 92

QY 62 ST-VVSTFTFTFTFOISTPTSPGDLAPFLAPYDTVIPNSAGNLLGLFPNLNLRNS 120

Db 93 TTGNVASEFTSFIIQAPNPATTADGLAFAPVDT--QPDGLGGMGLGIF-----KDG 144

QY 121 TTSKETTIDVNAASNNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWQNGKT 180

Db 145 SYNK-----SNOIVAVEFDTFSTNIHF-DPKGRHMGINVASIVSKTVPNNWTNGEV 194

QY 181 ATAHTSYNSAKRLSVTTFPGKGAVLSHD-VELTQVLQWIRVGFSASTGLE 233

Db 195 ANVFISYEASTKSLNASLVYPSLSTFIIHAIYDVKDLVPEWVRFGFSATTGID 248

RESULT 13

S62691

agglutinin II precursor - black locust

N:Alternate names: lectin

C:Species: Robinia pseudoacacia (black locust)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S62691; S62686

R:van Damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.

Plant Mol. Biol. 29, 1197-1210, 1995

A:Title: The seed lectins of black locust (Robinia pseudoacacia) are encoded by two genes

A:Reference number: S62685; MUID:96191285; PMID:8616218

A:Accession: S62691

A:Molecule type: mRNA

A:Residues: 1-285 <VAV>

A:Cross-references: EMBL:U24250

A:Experimental source: seed

A:Accession: S62686

A:Molecule type: protein

A:Residues: 32-50 <VAV>

C:Superfamily: plant lectin

C:Keywords: glycoprotein; homotetramer; lectin

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-285/Product: agglutinin II #status experimental <MAT>

Query Match

Best Local Similarity 36.7%; Score 444; DB 2; Length 285;

Matches 100; Conservative 46; Mismatches 56; Indels 22; Gaps 8;

QY 3 SVSFTTTFKDSQDKLMFQGH-TISSNVQLTKLDSNGNPVSTSGRVLVSAPLRLWES 61

Db 34 SLSEFPKFAFPNQPYLIFQDALVTSTGVLQLTNV-VNGVPPRRSIGRALYAAFFQIWDN 92

QY 62 ST-VVSTFTFTFTFOISTPTSPGDLAPFLAPYDTVIPNSAGNLLGLFPNLNLRNS 120

Db 93 TTGNVASEFTSFIIQAPNPATTADGLAFAPVDT--QPDGLGGMGLGIF-----KNG 144

QY 121 TTSKETTIDVNAASNNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVAVDWQNGKT 180

Db 145 YFNK-----SNOIVAVEFDTFSTNRH-WDPTGRHMGINVASIVSKTVPNNWANGEV 194

QY 181 ATAHTSYNSAKRLSVTTFPGKGAVLSHD-VELTQVLQWIRVGFSASTGLE 233

Db 195 ANVFISYEASTKSLNASLVYPSLSTFIIHAIYDVKDLVPEWVRFGFSATTGID 248

RESULT 14

JQ1981

lectin II - Scotch broom

N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II

C:Species: Cytisus scoparius (Scotch broom)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999

C:Accession: JQ1981

R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.

J. Biochem. 112, 366-375, 1992

A:Title: The primary structure of the Cytisus scoparius seed lectin and a carbohydrate-h

A:Reference number: JQ1981; MUID:93054441; PMID:1429525

A:Accession: JQ1981

A:Molecule type: protein

A:Residues: 1-249 <KON>

A:Experimental source: seed

C:Superfamily: plant lectin

C:Keywords: lectin

Query Match 36.4%; Score 439.5; DB 2; Length 249;

Best Local Similarity 41.7%; Pred. No. 3.2e-29;

Matches 101; Conservative 46; Mismatches 66; Indels 29; Gaps 9;

QY 1 AQSVSFTFTKFDQKDLMOGH--TISSSNVIOITKLDNSGNPVSTSVGRVLYSAPLR 58

DB 1 SEELSFSTFKTKDQKNLLEORDALITPTCKLQLTVE--NGKPAAYSILGRALYSTPIHI 59

QY 59 WESST-VVSTFTFTFTFOIS---TPYTSPPGDGLAFFLAPYDVTIPPNSAGNLLGLFPNL 114

DB 60 WDKSTGDEASFPATFESFVISPAPNPSTAAT--DGLAFFLAPADT--QPQSAGGYLGLF--- 113

QY 115 NALRNSITSKETTTIDVNAASNNVAVEDTYPNDNIGDPYRKHIGIDVNSIRSKATVAMD 174

DB 114 -----EKDSYN---SSNQIVAVEEDTYINSAMDPTQNPHTGIDVNTIKSKVSSWG 162

QY 175 WONGKTATAHISYNSASKRLSVTTFYPGGK-----AVLSHSDVELTQVLPOWIRVGFSSAS 229

DB 163 FKNGVATVLITYQPSKSLVASLVYPSGQTSKTSYIISANVDLKATVPWVRIGFSAT 222

QY 230 TG 231

DB 223 TG 224

RESULT 15

S51831

phytohemagglutinin L precursor - Phaseolus acutifolius

N:Alternate names: PHA-L protein

C:Species: Phaseolus acutifolius

C:Date: 14-Jul-1995 #sequence_revision 12-Apr-1996 #text_change 20-Aug-1999

C:Accession: S51831

R:Mirkov, T.E.; Wahlstrom, J.M.; Haglwar, K.; Finardi-Filho, F.; Kjemtrup, S.; Christee

Plant Mol. Biol. 26, 1103-1113, 1994

A:Title: Evolutionary relationships among proteins in the phytohemagglutinin-arcelin-alf

A:Reference number: S51827; MUID:95111094; PMID:7811969

A:Accession: S51831

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-276 <MIR>

A:Cross-references: EMBL:U10416; NID:9500744; PIDN:AAA82181.1; PID:9500745

C:Superfamily: plant lectin

F:1-24/Domain: signal sequence #status predicted <SIG>

F:25-276/Product: phytohemagglutinin L #status predicted <MAT>

Query Match 36.3%; Score 438.5; DB 2; Length 276;

Best Local Similarity 41.4%; Pred. No. 4.4e-29;

Matches 98; Conservative 44; Mismatches 70; Indels 25; Gaps 7;

QY 1 AQSVSFTFTKFDQKDLMOGH--TISSSNVIOITKLDNSGNPVSTSVGRVLYSAPLR 59

DB 25 ANDISFNQRF--NETNIILOGDASVSSGQLRLNLDNGEPTLSSIGRAFYSTPIQIW 82

QY 60 ESST-VVSTFTFTFTFOISTPYTPSGDGLAFFLAPYDVTIPPNSAGNLLGLFPNLALR 118

DB 83 DSTTGAVASFPATFTFNIRVPNNAGPADGLAFALVPVGS--KPKDRGGLGLF----- 133

QY 119 NSTTSKETTIDVNAASNNVAVEDTYPNDNIGDPYRKHIGIDVNSIRSKATVAMDQWQ 178

DB 134-----DGSDSRATVAVEFTLYNRD--WDRPRHIGIDVNSIKITTPWDFGOG 182

QY 179 KTAUHAHSYNSASKRLSVTTFYPGGK--AVLSHSDVELTQVLPOWIRVGFSSATGLEK 234

DB 183 EDAEVLITYDSSSTKLLVASLVYPSQKTSFTVSDTVLKSVLPEWVRVGFSSATSGITK 239

Search completed: February 26, 2003, 16:51:34

Job time : 19.5281 secs

result No.	Query No.	Score	Match %	Length	DB	ID	Description	
1	2	666	55.1	240	1	LEC_BOWMI	P42088	bowringia m
2	2	646.5	53.5	290	1	CONA_CANGL	P14894	canavalia c
3	3	642.5	53.1	290	1	CONA_CANEN	P02866	canavalia c
4	4	629.5	52.1	237	1	LECA_DOLIA	P38662	dolichos la
5	5	593.5	49.1	290	1	LEC2_CLALU	Q39529	cladrastis
6	6	566	46.8	293	1	LEC1_CLALU	Q39528	cladrastis
7	7	477.5	39.5	290	1	LEC3_CLALU	Q39527	cladrastis
8	8	469.5	38.8	286	1	LEC2_ROBPS	Q42372	robinia pse
9	9	468.5	38.4	240	1	LEC3_VATMA	P81371	vatairea ma
10	10	468	38.7	270	1	LEC3_SOPJA	P93538	sophora japa
11	11	455	37.6	292	1	LEC3_SOPJA	P93535	sophora japa
12	12	447	37.0	273	1	PHAM_SOPAV	P15231	phaseolus v
13	13	446	36.9	285	1	LEC1_SOPJA	Q41159	robinia pse
14	14	445	36.8	285	1	LCB1_ROBPS	Q41162	robinia pse
15	15	444	36.7	285	1	LCB1_ROBPS	Q41161	robinia pse
16	16	444	36.7	285	1	LEC2_ROBPS	P29257	cytiscus sco
17	17	443	36.6	248	1	LEC2_CYTSC	Q01807	medicago tr
18	18	441.5	36.5	280	1	LEC2_MEDTR	P24146	griffonia s
19	19	437	36.1	285	1	LEC3_GRISI	P05046	glycine max
20	20	431.5	35.7	277	1	LEC3_SOBYN	P05046	glycine max
21	21	420.5	34.8	249	1	LEC1_MEDTR	Q01806	medicago tr
22	22	420.5	34.8	272	1	LEC2_ULEEU	P23973	ulex europe
23	23	418.5	34.6	275	1	PHAE_PHAVU	P05087	phaseolus v
24	24	413	34.2	250	1	LEC1_PHAVU	P05088	phaseolus v
25	25	411.5	34.0	275	1	LEC1_LABAL	P23558	laburnum al
26	26	405.5	33.5	233	1	LEC3_PEA	P02867	pisum sativ
27	27	404	33.4	281	1	LEC_VICFA	P02871	vicia faba
28	28	402.5	33.3	236	1	LEC_ERYCO	P16404	erythrina c
29	29	402	33.3	275	1	LEC_ONOVO	P28874	onobrychis b
30	30	379	31.3	290	1	LEC1_DOLBI	P05045	dolichos bi
31	31	374	30.9	275	1	LEC_BAUPU	P16030	bauhinia pu
32	32	368	30.4	273	1	LEC5_DOLBI	P19588	dolichos bi
33	33	366.5	30.3	242	1	LEC3_ARAHY	P02872	arachis hyp
						LEC1_FSOYE	O34313	psophocarpu

[illegible]

RX MEDLINE=75095622; PubMed=1112813;
 RT Wang J.L., Cunningham B.A., Waxdal M.J., Edelman G.M.;
 RA "The covalent and three-dimensional structural of concanavalin A. I.
 RL Amino acid sequence of cyanogen bromide fragments F1 and F2.";
 RN J. Biol. Chem. 250:1490-1502(1975).
 [3]
 RP SEQUENCE OF 164-281.
 RX MEDLINE=75095623; PubMed=1112814;
 RA Cunningham B.A., Wang J.L., Waxdal M.J., Edelman G.M.;
 RT "The covalent and three-dimensional structure of concanavalin A. II.
 RL Amino acid sequence of cyanogen bromide fragment F3.";
 RN J. Biol. Chem. 250:1503-1512(1975).
 [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=73053316; PubMed=4638345;
 RA Hardman K.D., Ainsworth C.F.;
 RT "Structure of concanavalin A at 2.4-A resolution.";
 RL Biochemistry 11:4910-4919(1972).
 [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=75095624; PubMed=1112815;
 RA Becker J.W., Reeke G.N. Jr., Wang J.L., Cunningham B.A., Edelman G.M.;
 RT "The covalent and three-dimensional structure of concanavalin A. III.
 RL Structure of the monomer and its interactions with metals and
 RT saccharides.";
 RN J. Biol. Chem. 250:1513-1524(1975).
 [6]
 RP X-RAY CRYSTALLOGRAPHY.
 RX MEDLINE=75095625; PubMed=1112816;
 RA Reeke G.N. Jr., Becker J.W., Edelman G.M.;
 RT "The covalent and three-dimensional structure of concanavalin A. IV.
 RL Atomic coordinates, hydrogen bonding, and quaternary structure.";
 RN J. Biol. Chem. 250:1525-1547(1975).
 [7]
 RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
 RA Parkin S., Rupp B., Hope H.;
 RT "Atomic resolution structure of concanavalin A at 120 K.";
 RL Acta Crystallogr. D 52:1161-1168(1996).
 [8]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RX MEDLINE=20347885; PubMed=10748006;
 RA Bouckaert J., Dewalle Y., Poortmans F., Wyns L., Loris R.;
 RT "The structural features of concanavalin A governing non-proline
 RT peptide isomerization.";
 RL J. Biol. Chem. 275:19778-19787(2000).
 CC -!- FUNCTION: D-mannose specific lectin.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY
 CC 30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER
 CC POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL
 CC SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS
 CC FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW
 CC PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
 CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
 CC SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOSUS LECTIN FAMILY.
 CC -!- DATABASE: NAME=worthington enzyme manual;
 CC WWW="http://www.worthington-biochem.com/manual/C/CONA.html"
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X01632; CAA25787.1;
 DR PIR; A03357; CVJBP.
 DR PIR; A03358; CVJBP.
 DR PDB; 1CNI; 30-SEP-83.
 DR PDB; 2CNA; 31-JUL-94.

DR PDB; 3CNA; 15-APR-91.
 DR PDB; 5CNA; 15-OCT-94.
 DR PDB; 1CON; 31-JAN-94.
 DR PDB; 2CTV; 31-OCT-93.
 DR PDB; 1SCR; 15-OCT-94.
 DR PDB; 1SCS; 15-OCT-94.
 DR PDB; 1APN; 03-APR-96.
 DR PDB; 1CES; 15-FEB-97.
 DR PDB; 1CJP; 15-OCT-97.
 DR PDB; 1CVN; 14-OCT-96.
 DR PDB; 1ENQ; 17-AUG-96.
 DR PDB; 1ENR; 16-FEB-99.
 DR PDB; 1ENS; 17-AUG-96.
 DR PDB; 1GIC; 20-AUG-97.
 DR PDB; 1JBC; 12-FEB-97.
 DR PDB; 1NLS; 26-NOV-97.
 DR PDB; 1ONA; 17-SEP-97.
 DR PDB; 1TEL; 24-JUN-98.
 DR PDB; 1VAL; 11-JAN-97.
 DR PDB; 1VAM; 11-JAN-97.
 DR PDB; 1VLN; 01-APR-97.
 DR PDB; 1BXH; 07-OCT-98.
 DR PDB; 2CAU; 25-NOV-98.
 DR PDB; 2CAV; 25-NOV-98.
 DR PDB; 1DQ0; 19-JAN-00.
 DR PDB; 1DQ1; 19-JAN-00.
 DR PDB; 1DQ2; 19-JAN-00.
 DR PDB; 1DQ4; 19-JAN-00.
 DR PDB; 1DQ5; 19-JAN-00.
 DR PDB; 1DQ6; 19-JAN-00.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 KW Lectin; Calcium; Manganese; Glycoprotein; Signal; 3D-structure.
 FT SIGNAL 1 29
 FT CHAIN 30 148 CONCAVALIN (SECOND PART).
 FT PROPEP 149 163 CONCAVALIN (FIRST PART).
 FT CHAIN 164 281
 FT PROPEP 282 290
 FT SITE 148 149 CLEAVAGE.
 FT SITE 163 164 CLEAVAGE.
 FT SITE 281 282 CLEAVAGE.
 FT METAL 171 171 MANGANESE.
 FT METAL 173 173 MANGANESE AND CALCIUM.
 FT METAL 175 175 CALCIUM.
 FT METAL 177 177 CALCIUM.
 FT METAL 182 182 MANGANESE AND CALCIUM.
 FT METAL 187 187 MANGANESE.
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
 FT CONFLICT 33 33 E -> Q (IN REF. 2).
 FT CONFLICT 35 35 N -> D (IN REF. 2).
 FT CONFLICT 62 62 E -> D (IN REF. 3).
 FT CONFLICT 66 66 R -> E (IN REF. 3).
 FT CONFLICT 98 98 V -> T (IN REF. 3).
 FT CONFLICT 100 101 AS -> SA (IN REF. 3).
 FT CONFLICT 107 107 T -> A (IN REF. 3).
 FT STRAND 167 173
 FT STRAND 178 180
 FT HELIX 181 181
 FT TURN 187 192
 FT STRAND 199 202
 FT TURN 207 208
 FT TURN 210 218
 FT TURN 219 222
 FT STRAND 223 229
 FT TURN 231 232
 FT STRAND 236 241

Query Match	46.8%;	Score 566;	DB 1; Length 293;
Best Local Similarity	51.1%;	Pred. No. 1.8e-39;	
Matches 119;	Conservative 40;	Mismatches 64;	Indels 10; Gaps 7;

QY	1	AQSVSFTFKFDSQDKDLMFO-GHTISSNVIOQLTKLDSNGNPNVSTSVGRVLVSAPLRW	59
DG	36	SDLSFTFNPPNSDDIFOKDASISNETLELTRISSSQPATSSVGRLATYPVRLW	95
	 : : : : : : : : : : . .	
QY	60	ESST-VVSTFFETTFQISTPYTSPGGDLGAFFLPAYDTVIPPSAGNLLGLFPNLNALR	118
DG	96	DKSTGRULASFRTTESFAITSP-TQPGDGFAFIAPPDDT--PGVGGGLLGNGFN-LR	151
		: : : : : : : : : : :	
QY	119	NSTYSKETTIDVNAASNNAVVEDTYPDNIGDPYRKHIGIDVNSIRSKATVAWDQWG	178
DG	152	NSSNN--GVAVNNQSAQIVAEFDITYNGQC-DPKYRHVGIDVNSITSLAYTQMOWMG	207
		: : : : : : : : : : :	
QY	179	KATAHISYNSASKPLSVTTTPCGKAYSLSHDVELTOVLPQWRVGFSASTG	231
DG	208	VKATAQSYNPASQKLTAVTSPNSTPLTVSLDIDLQTVLPEWRVGFSASTG	260
		: : : : : : : : : :	

RESULT 7

LECR_CLALU	STANDARD:	PRT:	290 AA.
ID -LECR_CLALU			
AC Q39527;			
DT 15-DEC-1998 (Rel. 37, Created)			
DT 15-DEC-1998 (Rel. 37, Last sequence update)			
DT 15-DEC-1998 (Rel. 37, Last annotation update)			
DE Lectin-related protein precursor (CLLRP) (LRPCL) (Fragment).			
OS Cladrestis lutea (Yellow wood).			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC Eucotids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.			
OX NCBI_TaxId=38412;			
RN [1]			
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-56.			
RC TISSUE=Bark;			

Query Match 38.8%; Score 469.5; DB 1; Length 286;
 Best Local Similarity 45.1%; Pred. No. 1.4e-31;
 Matches 106; Conservative 42; Mismatches 66; Indels 21; Gaps 8;

QY 3 SVSFTFTKEDSKDQKDLMFQGH-T-ISSNVITQLTKLDSNGNPVSTSVGRVLSYAPLRWES 61
 DB 34 SLSEFPKFKHQPDILFQSDALVTSKGLVQLTV-NDGRPVYDSIGRVLYAAPFQIWD 92
 QY 62 ST-VVSTFTFTFQISTPTYPSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLAIRNS 120
 DB 93 TTGNVASFVTSFIIKAPNEGTADCLFLAPVGST-QPLKGGGLGLF-----KDE 145

QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVANDWONGKT 180
 DB 146 SYNK-----SNOIVAVEFDTPNR-VAMDPNGLHMGIDVNSIOSRVTRMDWANGEV 195

QY 181 ATAHISYNSAKRLSVTTFYP-GKAVLSLSDHVELTQVLPQWIRVGFSASTGLEK 234
 DB 196 ANVFIYSYEASTKSLTASLVPSLEKSFILSAIVDLKKVLPWVRVVGFTATTGLSE 250

RESULT 9
 LECS_VATWA
 ID LECS_VATWA STANDARD; PRT; 240 AA.
 AC P81371;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Seed lectin (VML).
 OS Vatairea macrocarpa.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids 1; Fabales; Fabaceae; Papilionoideae; Dalbergiaceae; Vataleae.
 OX NCBI_TaxID=77050;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=98218569; PubMed=9559667;
 RA Calvete J.J., Santos C.F., Mann K., Grangeiro T.B., Nimtz M.,
 Urbanke C., Sousa-Cavada B.;
 RT "Amino acid sequence, glycan structure, and proteolytic processing of
 the lectin of Vatairea macrocarpa seeds.";
 RL FEBS Lett. 425:286-292(1998).
 CC -!- FUNCTION: LECTIN THAT BINDS GALACTOSE.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- TISSUE SPECIFICITY: SEED.
 CC -!- PTM: PARTIALLY N-GLYCOSYLATED AT POSITIONS 111 AND 183 WITH
 THE HEPTASACCHARIDE [(BETA-Xylosyl-1,2)(ALPHA-MANNOSYL-1,6)(ALPHA-
 MANNOsyl-1,3)]BETA-MANOSYL-1,4-GLCNAC-BETA-1,4-GLCNAC-BETA-1,4
 [ALPHA-FUCOSYL-1,3]GLCNAC. A SMALL PROPORTION OF ALPHA CHAINS ARE
 PROTEOLYTICALLY CLEAVED AT 114-115 INTO GAMMA AND BETA CHAINS.
 CC THIS IS PROBABLY DEPENDENT ON THE DEGLYCOSYLATION OF N-111.
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
 AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
 SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES (BY
 SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 DR HSP; P19588; 1LUL.
 DR GlycosuitedB; P81371; -.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 KW Lectin; Calcium; Manganese; Glycoprotein.
 FT CHAIN 1 240 SEED LECTIN ALPHA CHAIN.
 FT CHAIN 1 114 SEED LECTIN GAMMA CHAIN.
 FT CHAIN 115 239 SEED LECTIN BETA CHAIN.
 FT METAL 123 123 MANGANESE (BY SIMILARITY).

FT METAL 125 125 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 129 129 CALCIUM (BY SIMILARITY).
 FT METAL 132 132 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 137 137 MANGANESE (BY SIMILARITY).
 FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .).
 FT VARIANT 117 117 I -> V.
 FT VARIANT 148 148 M -> K.
 FT VARIANT 154 154 G -> A.
 FT VARIANT 168 168 E -> Q.
 FT UNSURE 239 240
 SQ SEQUENCE 240 AA; 26197 MW; C17DF6B2568C65C1 CRC64;

Query Match 38.8%; Score 468.5; DB 1; Length 240;
 Best Local Similarity 42.6%; Pred. No. 1.4e-31;
 Matches 101; Conservative 43; Mismatches 68; Indels 25; Gaps 7;

QY 1 AOSVFTFTKEDSKDQKDLMFQGH-T-ISSNVITQLTKLDSNGNPVSTSVGRVLSYAPLRW 59
 DB 1 SEVVSFSTKPNPNPKDIILOGLDALVTSKGLQTKV-KDGKPDVDSHSLGRALYAAPIH 59

QY 60 ESST-VVSTFTFTFQISTPTYPSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLAIR 118
 DB 60 DDSTDVASEATSFVVEAPDESKTAGIAFFLAPDPT--QPKDGGFGLFENDSN--- 114

QY 119 NSTTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVANDWONG 178
 DB 115 -----KSIQTVAVEFDTF--SNWDPARHIGINVNIESMKYKVGWENG 158

QY 179 KTATAHISYNSAKRLSVTTFYP-GKAVLSLSDHVELTQVLPQWIRVGFSASTGLEK 234
 DB 159 KVANVVISYEASTKSLTASLVPSNATSVANSVNDLKSALPEWVRVGFSASTGLSR 215

RESULT 10
 LECS_SOPJA
 ID LECS_SOPJA STANDARD; PRT; 270 AA.
 AC P93538;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Bark lectin precursor (LECSJABG) (fragment).
 OS Sophora japonica (Japanese pagoda tree).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OX NCBI_TaxID=3897;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bark;
 RX MEDLINE=97201486; PubMed=9049272;
 RA van Damme E.J., Barre A., Rouge P., Peumans W.J.;
 RT "Molecular cloning of the bark and seed lectins from the Japanese
 pagoda tree (Sophora japonica).";
 RL Plant Mol. Biol. 33:523-536(1997).
 CC -!- FUNCTION: GALNAC-SPECIFIC LECTIN.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.
 CC
 CC EMBL; U63014; AAB51458.1; -.
 DR HSP; P19588; 1LUL.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.

```

DR ProDom: PD000711; LECTIN_LEG_B; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR LECTIN; Calcium; Manganese; Glycoprotein; Signal.
KW NON_TER 1
FT SIGNAL 1
FT CHAIN 1
FT METAL 16
FT METAL 141
FT METAL 143
FT METAL 145
FT METAL 147
FT METAL 150
FT METAL 155
FT METAL 155
FT CARBOHYD 60
FT CARBOHYD 76
FT CARBOHYD 127
FT CARBOHYD 201
FT CARBOHYD 221
SQ SEQUENCE 270 AA; 29314 MW; 1FD655A2C4E550B3 CRC64;

Query Match 38.7%; Score 468; DB 1; Length 270;
Best Local Similarity 43.8%; Pred. No. 1.8e-31;
Matches 103; Conservative 47; Mismatches 63; Indels 22; Gaps 8;

QY 1 AQSVSFTTFKDSOKDLMEFGHT-ISSNNVIQLKLDNGNPVSTSVGRVLYSAPLRW 59
DB 16 AEILSFSFPKFSVNGEDLLQGDALVSSGEGELQLTVE-NGVPVWNSTGRALYAPVHIW 74
QY 60 ESST-VVSTFTTFTFOISTPTSPPGDGLAFPLAPDYTVIPNSAGNIIGLFPNLNLR 118
DB 75 DNSTGRVASFATSEFVVKAPVASKSADGIAFLAPLNQI-HGAGGGLYGLF----- 126
QY 119 NSTSKETITIDVNAASNNVAVEFTYDNDNLGDPYRKHIGIDVNSIRKATVANDQNG 178
DB 127 NSSS-----YSSYQIVAVEFTDTH--TNAMDPNTRHIGIDVNSIRKATVANDQNG 176
QY 179 KATAHISYNSAKRLSVTFYPGGK-AVSLSHDVELTOVLPOWIRVGFSASTGL 232
DB 177 EVANVLLTYQAATEMLTSLYTPSNQTSYLSAAYDLKSLPFWVRVGTATGL 231

RESULT 11
LECS_SOPJA STANDARD; PRT; 292 AA.
AC P93535;
AT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Seed lectin precursor (LECSJASG).
OS Sophora japonica (Japanese pagoda tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.
OX NCBI_TaxID=3897;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=97201486; PubMed=9049272;
RA van Damme E.J., Barre A., Rouge P., Peumans W.J.;
RT "Molecular cloning of the bark and seed lectins from the Japanese
RL pagoda tree (Sophora japonica).";
RL Plant Mol. Biol. 33:523-536(1997).
CC -!- FUNCTION: MANNANOSE/GLUCOSE-SPECIFIC LECTIN.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or *send an email to license@isb-sib.ch).
CC -----
CC EMBL; U63011; AAB51441.1; -.

```

```

DR HSSP; P19588; ILUL.
DR GlycoSuiteDB; P93535; -.
DR InterPro; IPR000985; LECTIN_LEG.
DR InterPro; IPR001220; LECTIN_LEG.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; LECTIN_LEG; 1.
DR ProDom; PD000711; LECTIN_LEG; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW LECTIN; Calcium; Manganese; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 37
FT METAL 163
FT METAL 163
FT METAL 165
FT METAL 167
FT METAL 169
FT METAL 172
FT METAL 177
FT METAL 177
FT CARBOHYD 82
FT CARBOHYD 154
FT CARBOHYD 186
FT CARBOHYD 186
SQ SEQUENCE 292 AA; 31656 MW; A7431C29117A503E CRC64;

Query Match 37.6%; Score 455; DB 1; Length 292;
Best Local Similarity 43.3%; Pred. No. 2.3e-30;
Matches 104; Conservative 41; Mismatches 63; Indels 32; Gaps 9;

QY 1 AQSVSFTTFKDSOKDLMEFGHT-ISSNNVIQLKLDNGNPVSTSVGRVLYSAPLRW 59
DB 38 AEILSFSFPKFSVNGEDLLQGDALVSSGEGELQLTVE-NGVPVWNSTGRALYAPVHIW 96
QY 60 ESST-VVSTFTTFTFOISTPTSPPGDGLAFPLAPDYTVIPNSAGNIIGLFPNLNLR 113
DB 97 DKSTGRVASFATSEFVVKAPVASKSADGIAFLA-----PPNNIQPGGHLGLFHS 150
QY 114 LNALRSTSKETITIDVNAASNNVAVEFTYDNDNLGDPYRKHIGIDVNSIRKATVAM 173
DB 151 SG--YNS-----SYQIIAVDFDTH--INAWDPNTRHIGIDVNSIRKATVAM 193
QY 174 DWONGKTATAHISYNSAKRLSVTFYPGGK-AVSLSHDVELTOVLPOWIRVGFSASTGL 232
DB 194 GWONGEVANVLLSYQAATEMLTSLYTPSNQTSYLSAAYDLKSLPFWVRVGTATGL 253

RESULT 12
PHAM_PHAVU STANDARD; PRT; 273 AA.
AC P15231;
AT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Leucoagglutinating phytohemagglutinin precursor (PHA-L).
GN PLEC2.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PINTO U1111; TISSUE=Leaf;
RA Voelker T.A., Staswick P., Chrispeels M.J.;
RT "Molecular analysis of two phytohemagglutinin genes and their
RL expression in Phaseolus vulgaris cv. Pinto, a lectin-deficient
RL cultivar of the bean.";
RL EMBO J. 5:3075-3082(1986).
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC COMBINATION IS THE ONLY ONE DEVOID OF AGGLUTINATION ACTIVITY.
 CC ISOFORM B4 DISPLAYS MAXIMAL AGGLUTINATION ACTIVITY.
 CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN SEED. LOWER LEVELS IN THE
 CC FLOWER, AND THE BARK OF THE ROOTS. NO EXPRESSION IN LEAF. THE
 CC LECTIN ACCUMULATES IN THE INNER BARK IN AUTUMN.
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U12782; AAA80181.1; -
 CC HSSP; P19588; ILUL.
 CC InterPro; IPR000985; Lectin_legA.
 CC InterPro; IPR001220; Lectin_legB.
 CC Pfam; PF00138; lectin_legA; 1.
 CC Pfam; PF00139; lectin_legB; 1.
 CC ProDom; PD000671; Lectin_legA; 1.
 CC ProDom; PD000711; Lectin_legB; 1.
 CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
 KW SIGNAL 1 31
 FT CHAIN 1 32 285 BARK AGGLUTININ I, POLYPEPTIDE A.
 FT METAL 156 158 MANGANESE (BY SIMILARITY).
 FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 160 160 CALCIUM (BY SIMILARITY).
 FT METAL 162 162 CALCIUM (BY SIMILARITY).
 FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 171 171 MANGANESE (BY SIMILARITY).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 285 AA; 30928 MW; 49382E50EEF27382 CRC64;

Query Match 36.9%; Score 446; DB 1; Length 285;
 Best Local Similarity 42.6%; Pred. No. 1.2e-29;
 Matches 100; Conservative 47; Mismatches 66; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKLMFQGH-TISSNVITQLTKLDSNGNPVSTVGRVLYSAPLRLWES 61
 DB 34 SLSFEPKFPAPNOPIYLFQDALVTSTGVQLQTNV-VNGVPSGKSLGRALYAAPFQIWD 92
 QY 62 ST-VVSTFTFTTFTTQISTPTSPGDLAFPLAPYDVIIPNSAGNLLGLFPNLRNS 120
 DB 93 TTGNVASFVTSFIIQAPNPTTADGLAFPLAPVD--QPLDVGGMGLGF-----KDG 144
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHIGIDVNSIRSKATVANDWQNGKT 180
 DB 145 YNKK-----SNOIVAVEFDTFSNIHF-DPKGRHMGINNVNSIVIKITVPWNWNGEV 194
 QY 181 ATAHISNSAKRLSVTTFTYPPGKAVKLSHD-VELTQVLQWIRVGFSASTGLEK 234
 DB 195 ANVFISYEASTKSLASLVYPSLETSPVHAIVDVKDVLPEWVRFGFSATTGIDK 249

RESULT 14
 LCBL_ROBPS STANDARD; PRT; 285 AA.
 ID LCBL_ROBPS
 AC Q41159;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Seed agglutinin I precursor (RPSAI) (LECRPASI).
 OS Robinia pseudoacacia (Black locust).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
 OX NCBI_TaxID=35938;
 RN [1]

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC COMBINATION IS THE ONLY ONE DEVOID OF AGGLUTINATION ACTIVITY.
 CC ISOFORM B4 DISPLAYS MAXIMAL AGGLUTINATION ACTIVITY.
 CC -!- TISSUE SPECIFICITY: STRONG EXPRESSION IN SEED. LOWER LEVELS IN THE
 CC FLOWER, AND THE BARK OF THE ROOTS. NO EXPRESSION IN LEAF. THE
 CC LECTIN ACCUMULATES IN THE INNER BARK IN AUTUMN.
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U12782; AAA80181.1; -
 CC HSSP; P19588; ILUL.
 CC InterPro; IPR000985; Lectin_legA.
 CC InterPro; IPR001220; Lectin_legB.
 CC Pfam; PF00138; lectin_legA; 1.
 CC Pfam; PF00139; lectin_legB; 1.
 CC ProDom; PD000671; Lectin_legA; 1.
 CC ProDom; PD000711; Lectin_legB; 1.
 CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
 KW SIGNAL 1 31
 FT CHAIN 1 32 285 BARK AGGLUTININ I, POLYPEPTIDE A.
 FT METAL 156 158 MANGANESE (BY SIMILARITY).
 FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 160 160 CALCIUM (BY SIMILARITY).
 FT METAL 162 162 CALCIUM (BY SIMILARITY).
 FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 171 171 MANGANESE (BY SIMILARITY).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 285 AA; 30928 MW; 49382E50EEF27382 CRC64;

Query Match 36.9%; Score 446; DB 1; Length 285;
 Best Local Similarity 42.6%; Pred. No. 1.2e-29;
 Matches 100; Conservative 47; Mismatches 66; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKLMFQGH-TISSNVITQLTKLDSNGNPVSTVGRVLYSAPLRLWES 61
 DB 34 SLSFEPKFPAPNOPIYLFQDALVTSTGVQLQTNV-VNGVPSGKSLGRALYAAPFQIWD 92
 QY 62 ST-VVSTFTFTTFTTQISTPTSPGDLAFPLAPYDVIIPNSAGNLLGLFPNLRNS 120
 DB 93 TTGNVASFVTSFIIQAPNPTTADGLAFPLAPVD--QPLDVGGMGLGF-----KDG 144
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDNIGDPYRKHIGIDVNSIRSKATVANDWQNGKT 180
 DB 145 YNKK-----SNOIVAVEFDTFSNIHF-DPKGRHMGINNVNSIVIKITVPWNWNGEV 194
 QY 181 ATAHISNSAKRLSVTTFTYPPGKAVKLSHD-VELTQVLQWIRVGFSASTGLEK 234
 DB 195 ANVFISYEASTKSLASLVYPSLETSPVHAIVDVKDVLPEWVRFGFSATTGIDK 249

RESULT 14
 LCBL_ROBPS STANDARD; PRT; 285 AA.
 ID LCBL_ROBPS
 AC Q41159;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Seed agglutinin I precursor (RPSAI) (LECRPASI).
 OS Robinia pseudoacacia (Black locust).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
 OX NCBI_TaxID=35938;
 RN [1]

```

RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-51.
RX MEDLINE=961191285; PubMed=8616218;
RA van Damme E.J.M., Barre A., Rouge P., van Leuven F., Peumans W.J.;
RT "The seed lectins of black locust (Robinia pseudoacacia) are encoded
RL by two genes which differ from the bark lectin genes.";
RN Plant Mol. Biol. 29:1197-1210(1995).
CC -!- FUNCTION: SEED LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SEED.
CC -!- PTM: THREE OF THE FOUR POTENTIAL GLYCOSYLATION SITES ARE OCCUPIED.
CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24250; AAC49272.1;
CC HSSP; P19588; ILUL.
CC InterPro: IPR000985; Lectin_legA.
CC InterPro: IPR001220; Lectin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 1.
CC ProDom; PD000671; Lectin_legA; 1.
CC ProDom; PD000711; Lectin_legB; 1.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 285 SEED AGGLUTININ I.
FT METAL 156 156 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 162 162 CALCIUM (BY SIMILARITY).
FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 171 171 MANGANESE (BY SIMILARITY).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 285 AA; 30943 MW; 6AE82CDC920224CE CRC64;

Query Match 36.8%; Score 445; DB 1; Length 285;
Best Local Similarity 41.9%; Pred. No. 1.5e-29;
Matches 98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TSSNNVIOLTKLDSNGNPVSTSVGRVLYSAPRLWES 61
DB 34 SLSPFPKFAFPNQLIFQRDALVTSTGVQLTNN-VNGVPPRRSISGRALYAAPFQIWDN 92

QY 62 ST-VVSTFTFTTFOISTPTSPPGDGLAFFLAPVDYVIPPNSAGNLLGLFPNLNLRNS 120
DB 93 TTGNVASFVTSFSFIQAPNATTADGLAFFLAPVDY-QPGDLGGMLGIF-----KDG 144

QY 121 TTSKETIDVNAASNNVAVFEDYIPDNIGDPYRKHGIDVNSRATKATVADWQNGKT 180
DB 145 SYNK-----SNQIVAVEFDTSNIHF-DPKGRHMGINNVNSIVSKTVPNWNTNGEV 194

QY 181 ATAHISYNSAKRLSVTTFYPGGKAVLSHD-VELTOVLQWIRVGFSASTGLE 233
DB 195 ANVFISYEASTKSLNASLYPSLSTFSFIHAIVDVKVLPWVRFVGFSAATTGID 248

RESULT 15
LCS2_ROBPS STANDARD; PRT; 285 AA.
AC Q41161.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

```

DE SEQUENCE FROM N.A., AND SEQUENCE OF 32-50.
OS Robinia pseudoacacia (Black locust).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
OX NCBI_TaxID=35938;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-50.
RX TISSUE=Seed;
CC MEDLINE=961191285; PubMed=8616218;
RA van Damme E.J.M., Barre A., Rouge P., van Leuven F., Peumans W.J.;
RT "The seed lectins of black locust (Robinia pseudoacacia) are encoded
RL by two genes which differ from the bark lectin genes.";
RN Plant Mol. Biol. 29:1197-1210(1995).
CC -!- FUNCTION: SEED LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: EXPRESSION IN SEED.
CC -!- PTM: MOSTLY FOUND IN NON-GLYCOSYLATED FORM.
CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24249; AAC49271.1;
CC HSSP; P19588; ILUL.
CC InterPro: IPR000985; Lectin_legA.
CC InterPro: IPR001220; Lectin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 1.
CC ProDom; PD000671; Lectin_legA; 1.
CC ProDom; PD000711; Lectin_legB; 1.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 285 SEED AGGLUTININ II.
FT METAL 156 156 MANGANESE (BY SIMILARITY).
FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 162 162 CALCIUM (BY SIMILARITY).
FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 171 171 MANGANESE (BY SIMILARITY).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 285 AA; 31021 MW; 2C0B3249620294DE CRC64;

Query Match 36.7%; Score 444; DB 1; Length 285;
Best Local Similarity 42.7%; Pred. No. 1.8e-29;
Matches 100; Conservative 46; Mismatches 56; Indels 22; Gaps 8;

QY 3 SVSFTFTKFDSDQKDLMFQGH-TSSNNVIOLTKLDSNGNPVSTSVGRVLYSAPRLWES 61
DB 34 SLSPFPKFAFPNQLIFQRDALVTSTGVQLTNN-VNGVPPRRSISGRALYAAPFQIWDN 92

QY 62 ST-VVSTFTFTTFOISTPTSPPGDGLAFFLAPVDYVIPPNSAGNLLGLFPNLNLRNS 120
DB 93 TTGNVASFVTSFSFIQAPNATTADGLAFFLAPVDY-QPLDLGGMLGIF-----KNG 144

QY 121 TTSKETIDVNAASNNVAVFEDYIPDNIGDPYRKHGIDVNSRATKATVADWQNGKT 180
DB 145 YFNK-----SNQIVAVEFDTSNRH-WDPTGRHMGINNVNSIVSKTVPNWNTNGEV 194

QY 181 ATAHISYNSAKRLSVTTFYPGGKAVLSHD-VELTOVLQWIRVGFSASTGLE 233
DB 195 ANVFISYEASTKSLNASLYPSLSTFSFIHAIVDVKVLPWVRFVGFSAATTGID 248

Search completed: February 26, 2003, 16:52:56
Job time : 6.7191 secs

```

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 ; Search time 22.2022 Seconds
(without alignments)
2171.628 Million cell updates/sec

Title: US-09-476-485A-8

Perfect score: 1209

Sequence: 1 AQSVSFTFTKDSQKDLMF.....TOVLQWIRVGFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	796.5	65.9	272	10 Q9ZTA9	Q9ZTA9 dolichos la
2	776.5	64.2	279	10 Q9M7M4	Q9M7M4 phaseolus v
3	649.5	53.7	290	10 O04672	O04672 canavalia b
4	643.5	53.2	290	10 Q947H0	Q947H0 canavalia e
5	547	45.2	266	10 P93536	P93536 sophora jap
6	544	45.0	284	10 Q9FVU9	Q9FVU9 sophora fla
7	532	44.0	293	10 P93537	P93537 sophora fla
8	488.5	40.4	254	10 Q43376	Q43376 arachis hyp
9	485.5	40.2	280	10 Q43374	Q43374 arachis hyp
10	480.5	39.7	282	10 P93247	P93247 maackia amu
11	480.5	39.7	286	10 P93248	P93248 maackia amu
12	477.5	39.5	254	10 Q43377	Q43377 arachis hyp
13	459	38.0	258	10 Q9FVF8	Q9FVF8 ulex europe
14	457	37.8	285	10 Q9ZNP6	Q9ZNP6 robinia pse
15	449.5	37.2	279	10 Q49899	Q49899 medicago sa
16	449	37.1	256	10 P93246	P93246 maackia amu

17	446	36.9	275	10 Q8RVY1	Q8RVY1 phaseolus v
18	444.5	36.8	273	10 Q8RVY4	Q8RVY4 phaseolus c
19	443.5	36.7	275	10 Q8RVH1	Q8RVH1 phaseolus v
20	438.5	36.3	276	10 Q40750	Q40750 phaseolus a
21	436	36.1	274	10 Q43628	Q43628 phaseolus v
22	434.5	35.9	275	10 Q8RVX5	Q8RVX5 phaseolus v
23	433.5	35.9	273	10 Q42411	Q42411 medicago sa
24	433.5	35.9	273	10 Q8RVH2	Q8RVH2 phaseolus v
25	427.5	35.4	275	10 Q8RW23	Q8RW23 phaseolus c
26	427.5	35.4	275	10 Q8RVX6	Q8RVX6 phaseolus v
27	427.5	35.4	275	10 Q8RVH3	Q8RVH3 phaseolus v
28	426.5	35.3	275	10 Q8RVY0	Q8RVY0 phaseolus v
29	424.5	35.1	278	10 Q9LED9	Q9LED9 phaseolus l
30	424.5	35.1	278	10 Q9LED8	Q9LED8 phaseolus l
31	421.5	34.9	251	10 Q8WIR7	Q8WIR7 griffonia s
32	421	34.8	285	10 Q9ZWP5	Q9ZWP5 robinia pse
33	418.5	34.6	247	10 Q8WIR6	Q8WIR6 griffonia s
34	417.5	34.5	251	10 Q93X48	Q93X48 lens ervoid
35	417.5	34.5	251	10 Q93WH6	Q93WH6 lens culina
36	417.5	34.5	273	10 Q8RVX9	Q8RVX9 phaseolus v
37	416.5	34.4	251	10 Q93X49	Q93X49 lens culina
38	415.5	34.4	251	10 Q93X41	Q93X41 lens odomet
39	415.5	34.4	251	10 Q8W4Y4	Q8W4Y4 lens lamott
40	415.5	34.4	255	10 Q40912	Q40912 phaseolus l
41	414.5	34.3	251	10 Q8VXF2	Q8VXF2 lens culina
42	414.5	34.3	268	10 P93458	P93458 phaseolus l
43	414	34.2	240	10 Q9SM56	Q9SM56 psophocarpu
44	413.5	34.2	251	10 Q93X50	Q93X50 lens culina
45	412.5	34.1	251	10 Q8RW33	Q8RW33 lathyrus sa

ALIGNMENTS

RESULT 1

Q9ZTA9 PRELIMINARY; PRT; 272 AA.
ID Q9ZTA9
AC Q9ZTA9;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mannose lectin.
GN FRIL.
OS Dolichos lab lab (Field bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=35936;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RX MEDLINE=99110944; PubMed=9892687;
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;
RT "CDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves
hematopoietic progenitors in suspension culture.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).
DR EMBL; AF067417; AADL0734.1;
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; Lectin_legA_1.
DR Pfam; PF00139; Lectin_legB_1.
DR ProDom; PD000671; Lectin_legB; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 65.9%; Score 796.5; DB 10; Length 272;
Best Local Similarity 66.7%; Pred. No. 6.7e-57;
Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

QY 1 AOSVSFTFTKFDSDOKLDMFOGHITSSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 60
 DB 9 AQSLSFSTFKFDNQEDLIIFQGHATSNNVQLTKLDSAGNPVSSAGRVLYSAPLRLWE 68
 QY 61 SSTVSTFTFTFQISTPTSPGDLGFLAPDYDVIPPNNSAGNLLGLFPNLNLRNS 120
 DB 69 DSAVLTSFTDIINFEISTPTSRITADGLAFIAPDSDVI--SYHGGLGLFPNANTLNNS 126
 QY 121 TTSKETTIDVNAASN--NVAVERTDTPNDNIGDPYRKHGIDVNSIRSKATVAMDWQNGKT 180
 DB 127 STSENQT-TTKAASSNVAVFEDLYLPDYGDPNKHIGIDVNSIRSKATVAKWQNGKI 185
 QY 181 ATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
 DB 186 ATAHISYNSKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDK 239

RESULT 2

Q9M7M4 PRELIMINARY; PRT; 279 AA.
 AC Q9M7M4;
 DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Mannose lectin FRIL (Fragment).
 OS Phaseolus vulgaris (Kidney bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
 OX NCBI_TaxID=3885;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
 RA Feldman M.;
 RT "A new lectin in red kidney bean called PvFRIL stimulates
 RT proliferation of NIH3T3 cells expressing the Flt3 receptor";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF121458; AAF28739.1; -.
 DR HSSP; P02866; IONA.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; lectin_legA; 1.
 DR ProDom; PD000711; lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 KW Lectin.
 FT NON_TER 1
 SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;

Query Match 64.2%; Score 776.5; DB 10; Length 279;
 Best Local Similarity 66.8%; Pred. No. 2.9e-55;
 Matches 157; Conservative 25; Mismatches 48; Indels 5; Gaps 4;

QY 1 AOSVSFTFTKFDSDOKLDMFOGHITSSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 60
 DB 1 AQSLSFSTFKFDLQKLIQFQGHATSNNVQLTKLDSNGPNVSGRVLYSAPFLWE 60
 QY 61 SSTVSTFTFTFQISTPTSPGDLGFLAPDYDVIPPNNSAGNLLGLFPNLNLRNS 120
 DB 61 NSMAVSFTFTFQISTPTSPHYYAAGGFAFLAPHDVTIPPNNSGKFLGLYSNV--FRNS 118
 QY 121 TTSKETTIDVNAASN--NVAVERTDTPNDNIGDPYRKHGIDVNSIRSKATVAMDWQNGK 179
 DB 119 PTSENQSGDVNTDS--RVAVERTDTPNANI--DPNVRHIGIDVNSIKSKETARWEQNGK 176

QY 180 TATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234

DB 177 TATAHISYNSAKRLSVTTFYPPGMEVVALSHDVELTQVLPQWIRVGFSASTGLEK 231

RESULT 3

004672 PRELIMINARY; PRT; 290 AA.
 AC 004672;
 DT 01-JUL-1997 (T-EMBLrel. 04, Created)
 DT 01-JUL-1997 (T-EMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Lectin (Fragment).
 OS Canavalia brasiliensis (Brazilian jack bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
 OX NCBI_TaxID=61861;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Grangeiro T.B.;
 RL Thesis (1996), Universidade Federal do Ceara, Fortaleza, Brazil.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97454283; PubMed=9310358;
 RA Grangeiro T.B., Schrieffer A., Calvete J.J., Raida M., Urbanke C.,
 RA Barral-Netto M., Cavada B.S.;
 RT "Molecular cloning and characterization of ConBr, the lectin of
 RT Canavalia brasiliensis seeds";
 RL Eur. J. Biochem. 248:43-48(1997).
 DR EMBL; Y13904; CAA74202.1; -.
 DR HSSP; P02866; IONA.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; lectin_legA; 1.
 DR ProDom; PD000711; lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 FT NON_TER 290
 SQ SEQUENCE 290 AA; 31408 MW; 0CE5B4BBC950AC92 CRC64;

Query Match 53.7%; Score 649.5; DB 10; Length 290;
 Best Local Similarity 54.5%; Pred. No. 6.2e-45;

Matches 128; Conservative 35; Mismatches 63; Indels 9; Gaps 4;

QY 3 SVSFTFTKFDSDOKLDMFOGH-TISSNNVQLTKLDSNGPNVSTSVGRVLYSAPLRLWE 61
 DB 36 ALHFENQFSKQKDLILQGDATGTGDNLELTRVSSNGSPQSSVGRALFYAPVHIWES 95
 QY 62 STVSTFTFTFQISTPTSPGDLGFLAPDYDVIPPNNSAGNLLGLFPNLNLRNST 121
 DB 96 SAVVASFEATFTFLIKSP-DSHPADGIAFTISNIDSSIPSGTGLLGLFPDANVRNS- 153
 QY 122 TSKETTIDVNAASN--NVAVERTDTPNDNIGDPYRKHGIDVNSIRSKATVAMDWQNGK 179
 DB 154 ----TIDFNAAYNADTIVAVELDTYPTNDIGDPSYPHIGIDIKSVRSKKTAKNMQNGK 209
 QY 180 TATAHISYNSAKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
 DB 210 VCTAHIIYNSVGRKLSAVSVPNGDSATVSDVDLDNVLPWVRVGLSASTGLYK 264

RESULT 4

Q947H0 PRELIMINARY; PRT; 290 AA.
 AC Q947H0;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Concanavalin A.
 GN CONA.
 OS Canavalia ensiformis (Jack bean) (Horse bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
 OX NCBI_TaxID=3823;
 RN [1]

DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW	Lectin; Glycoprotein; Signal; Mannose-binding; Multigene family.
FT	SIGNAL 1 35
FT	CHAIN 36 162 BARK LECTIN I, A1 SUBUNIT.
FT	CHAIN 163 293 BARK LECTIN I, B1 SUBUNIT.
FT	CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 145 145 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 152 152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 213 213 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CONFLICT 57 R > S (IN REF. 2).
SEQ	SEQUENCE 293 AA; 32321 MW; 3D2F19IAD63F1986 CRC64;
Query Match 44.0%; Score 532; DB 10; Length 293;	
Best Local Similarity 47.5%; Pred. No. 2.le-35;	
Matches 112; Conservative 42; Mismatches 66; Indels 16; Gaps	
QY	1 AQSVFTTFRKFDSDOKDLFMFGHTISSNNVQLTKLDSNGNPSTSGRVLSAPRLWE 60
DB	36 SDLSLFTVENFPNPDILLORDASITSNETLLTRTSNGKPQKGSVGRALYYAPVRLWD 95
QY	61 SST-VVSFTFTTFQISTPYTSPGDLGAFFLAFLPYTPVPNSAGNLLGLFPNLNALRN 11
DB	96 KSTGRSLASFETSFSSVTISP-TTDPGGGIAFFAIAPPDTT--PGYTGGLLGLF-----N 14
QY	120 STTKSETTIIDVAASN---VVAVEPTDPNDNIGDPYRKHGIGDVNSIRSKATVAWDW 17
DB	146 SSTVQSNSDHGVAFHNSLPQIVAEVDITYNGG-RPNYNHVGIDVNSIKSVSTTKWTW 20
QY	176 QNGKTATAHISYNSASKRSLSYTYFPGKAVSLSHDVELTVLPQWIRVGFSASTG 231
DB	205 RNGVEATANISYNPVSQRLTAVSSYPNSEPI TVHYDIDLKTLVLPKWVRGFSASTG 260
RESULT 8	
ID	Q43376 PRELIMINARY; PRT; 254 AA.
AC	O43376;
DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Mannose/glucose-binding lectin precursor (fragment).
GN	LEC.
OS	Arachis hypogaea (peanut).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid1; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC	Arachis.
NCBI_TaxID=3818;	
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV. SELLIE; TISSUE=SEED;
RA	Law J.J.;
RT	"Cloning and expression of cDNA for mannose-binding lectin from
RT	peanut.";
RL	Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR	EMBL; U22472; AAA74575.1; ..
DR	HSSP; P02867; ZBQP.
DR	InterPro; IPR000985; Lectin_legA.
DR	InterPro; IPR001220; Lectin_legB.
DR	Pfam; PF00138; lectin_legA; 1.
DR	Pfam; PF00139; lectin_legB; 1.
DR	ProDom; PD000671; Lectin_legA; 1.
DR	ProDom; PD000711; Lectin_legB; 1.
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW	Lectin; Signal.
FT	NON_TER 1 1
FT	SIGNAL <1 2
SEQ	SEQUENCE 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;
Query Match 40.4%; Score 488.5; DB 10; Length 254;	
Best Local Similarity 46.6%; Pred. No. 6e-32;	
Matches 109; Conservative 42; Mismatches 64; Indels 19; Gaps	

```

QY 3 SVSFTTKFD-SDQKDLMOFGH-TISSNNVQLTKLDSNGNPVSTSVGRVLSAPLRWE 60
Db 3 SLFSFNKFEQDDERNILQGDATFSASKGQLTKVDANGTPAKSTVGRVLSHSTQVRLWE 62
QY 61 SST-VVSTFTTFTTQISTPTSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLNLRN 119
Db 63 KSTNLTNFOAQSFVLIKSP-NDIGADGIAFFIAAPDSQIPKNSAGGTLGLFDPTQA--- 118
QY 120 STTSKETTIDVNAASNNVAVFEDT-YPNDNIG-DPYRKHGIGIDVNSTRSKATVANDWQW 177
Db 119 -----QNPSANQVLAVEEDFTFAQDSNGWDPNYQHIGIDVNSIKSAATTKWERN 168
QY 178 GKATAHISYNSAKRLSVTFYPPGKRAVSLSHDVELTQVLPQWIRVGFSASTG 231
Db 169 GQTLNVLVTIDANSKNLQVTASYPDGQRYQVSYVVDLRDLHPWGRVGFSAASG 222

RESULT 9
ID Q43374 PRELIMINARY; PRT; 280 AA.
AC Q43374;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Mannose/glucose-binding lectin precursor.
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CV. SELLIE; TISSUE=MODULE;
RA Law I.J.;
RT "Cloning and expression of cDNA for mannose-binding lectin from
RL peanut";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22469; AAA74572.1; -.
DR HSSP; U2866; IDQ2.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; Lectin_legA; 1.
DR Pfam; PF00139; Lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT CHAIN 1 26 MANNOSE/GLUCOSE-BINDING LECTIN.
FT SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;

Query Match 40.2%; Score 485.5; DB 10; Length 280;
Best Local Similarity 45.7%; Pred. No. 1.2e-31;
Matches 107; Conservative 47; Mismatches 61; Indels 19; Gaps 8;

QY 3 SVSFTTKFD-SDQKDLMOFGH-TISSNNVQLTKLDSNGNPVSTSVGRVLSAPLRWE 60
Db 29 SLFSFNKFEQDDERNILQGDATFSASKGQLTKVDANGTPAKSTVGRVLSHSTQVRLWE 88
QY 61 SST-VVSTFTTFTTQISTPTSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLNLRN 119
Db 89 KSTNLTNFOAQSFVLIKSP-IDN-GADGIAFFIAAPDSQIPKNSAGGTLGLF-----D 140
QY 120 STTSKETTIDVNAASNNVAVFEDT-YPNDNIG-DPYRKHGIGIDVNSTRSKATVANDWQW 177
Db 141 PSTAQ-----NPSANQVLAVEEDFTFAQDSNGWDPNYQHIGIDVNSIKSAATTKWERN 194
QY 178 GKATAHISYNSAKRLSVTFYPPGKRAVSLSHDVELTQVLPQWIRVGFSASTG 231
Db 195 GQTLNVLVSIDANSKNLQVTASYPDGQRYQVSYNVDLRDLHPWGRVGFSAASG 248

```

```

RESULT 10
P93247 PRELIMINARY; PRT; 282 AA.
ID P93247;
AC P93247;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE BARK LEUCOAGGLUTININ I precursor (MALBI) (Fragment).
GN LECMALBI.
OS Maackia amurensis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.
OX NCBI_TaxID=37501;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=BARK;
RX MEDLINE=97390228; PubMed=9249142;
RA van Damme E.J.M.; Van Leuven F.; Peumans W.J.;
RT "Isolation, characterization and molecular cloning of the bark lectins
RT from Maackia amurensis.";
RL Glycoconj. J. 14:449-456(1997).
CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
CC -!- PTM: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR EMBL; U65009; AAB39933.1; -.
DR HSSP; P19588; LLUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR InterPro; IPR000531; TonB_boxC.
DR Pfam; PF00138; Lectin_legA; 1.
DR Pfam; PF00139; Lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; UNKNOWN_1.
DR PROSITE; PS00430; TONE_DEPENDENT_REC_1; UNKNOWN_1.
KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
FT NON_TER 1 24 BY SIMILARITY.
FT SIGNAL <1 24 BARK LEUCOAGGLUTININ I.
FT CHAIN 25 282 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 282 AA; 30879 MW; 45B3F714E1DB957A CRC64;

Query Match 39.7%; Score 480.5; DB 10; Length 282;
Best Local Similarity 45.3%; Pred. No. 3.1e-31;
Matches 107; Conservative 41; Mismatches 67; Indels 21; Gaps 7;

QY 1 AQSVSFTTKFDSDQKDLMOFGH-TISSNNVQLTKLDSNGNPVSTSVGRVLSAPLRWE 59
Db 25 SDELSFTINNPLNEADLLFOGEASVSTGVQLTRVE-NGQPKQKYSVGRALYAAPVRIW 83
QY 60 ESST-VVSTFTTFTTQISTPTSPGDLGAFFLAPYDTVIPPNSAGNLLGLFPNLNLRN 118
Db 84 DNTTGVASFSTFTFYVKAPEPNTSNGLAFFLAPDSQIPTGVTYGLGFNN----- 138
QY 119 NSTTSKETTIDVNAASNNVAVFEDTY--PNDNIGDPYRKHGIGIDVNSTRSKATVANDWQ 176
Db 139 ---TSDD-----SSNOI VAVEEDTFLHKYNPWPDPNVRHIGIDVNGIDSITKTQWDWI 188
QY 177 NGKTATAHISYNSAKRLSVTFYPPGGR-AVSLSHDVELTQVLPQWIRVGFSASTG 231

```


Db 185 PWSLVNKKANVVGAINVLSVDVEYPLVHRHTLSHVPLKDWVPEWVRIGFSAATG 244

QY 232 LE 233

Db 245 DE 246

Search completed: February 26, 2003, 16:54:27
Job time : 23.2022 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 28.6292 Seconds
(without alignments)
1089.120 Million cell updates/sec

Title: US-09-476-485A-8

Perfect score: 1209

Sequence: 1 AQSVSFTTKFDSQDKLMF.....TQVLPOWIRGVFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1193	98.7	234	AA662901	Amino acid sequenc
2	796.5	65.9	264	AA887973	A lectin derived p
3	796.5	65.9	264	AA662890	Amino acid sequenc
4	796.5	65.9	286	AA662894	Alpha-amylase inhi
5	776.5	64.2	303	AA662898	Amino acid sequenc
6	776.5	64.2	290	AAV58736	Jack bean concanav
7	638.5	52.8	290	AAV58737	Jack bean concanav
8	537.5	44.5	237	AA74765	Legume concanavall
9	457	37.8	285	20 AAY06812	R. pseudoacacia le
10	446	36.9	285	20 AAY06811	R. pseudoacacia le

11	445	36.8	285	20 AAY06814	Amino acid sequenc
12	444	36.7	285	20 AAY06815	Amino acid sequenc
13	411.5	34.0	275	21 AAY58738	Pea lectin. Pisum
14	379	31.3	242	15 AAR45911	Bauhinia purpurea l
15	379	31.3	262	20 AAY27265	B. purpurea native
16	371	30.7	242	15 AAR45912	Bauhinia purpurea l
17	368.5	30.5	239	16 AAR74767	Pea lectin, lile.
18	367	30.4	236	16 AAR64807	PNA lectin subunit
19	361.5	29.9	228	16 AAR74766	Pea lectin, 2lth.
20	341	28.2	274	20 AAY06813	R. pseudoacacia le
21	269	22.2	282	21 AAB25452	Pinus radiata cell
22	254	21.0	632	21 AAB25109	Pinus radiata cell
23	249	20.6	669	23 ABB93142	Herbicidally activ
24	244	20.2	265	10 AAP91967	Sequence of arceli
25	238.5	19.7	681	23 ABB93833	Herbicidally activ
26	236	19.5	652	23 ABB93374	Herbicidally activ
27	230	19.0	272	23 ABB91343	Herbicidally activ
28	223.5	18.5	244	10 AAP93640	Sequence of a Phas
29	223	18.4	677	23 ABB92654	Herbicidally activ
30	219	18.1	675	23 ABB92029	Herbicidally activ
31	212.5	17.6	674	23 ABB93878	Herbicidally activ
32	211	17.5	691	23 ABB93881	Herbicidally activ
33	208.5	17.2	715	23 ABB92649	Herbicidally activ
34	207	17.1	283	23 ABB91344	Herbicidally activ
35	206	17.0	627	23 ABB91918	Herbicidally activ
36	202	16.7	623	23 ABB91919	Herbicidally activ
37	199	16.5	685	23 ABB93319	Herbicidally activ
38	196.5	16.3	666	23 ABB91530	Herbicidally activ
39	196.5	16.3	693	23 ABB92247	Herbicidally activ
40	194.5	16.1	688	23 ABB93318	Herbicidally activ
41	194.5	16.1	711	23 ABB93337	Herbicidally activ
42	192	15.9	656	23 ABB91011	Herbicidally activ
43	188	15.6	661	23 ABB92738	Herbicidally activ
44	185.5	15.3	649	23 ABB93134	Herbicidally activ
45	181.5	15.0	664	23 ABB92106	Herbicidally activ

ALIGNMENTS

RESULT 1

AA662901
ID AA662901 standard; Protein; 234 AA.

XX AA662901;

XX 17-SEP-2001 (first entry)

XX Amino acid sequence of a partial FRIL polypeptide.

XX FRIL; FIK2/Flt3 tyrosine kinase receptor-interacting lectin;
KW progenitor cell preservation factor; radiotherapy; chemotherapy;
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Sphenostylis stenocarpa.

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume progenitor cell preservation factors for in vivo or ex vivo
PT preservation of hematopoietic progenitor cells and as therapeutics for

PT alleviating/reducing progenitor cell-depleting activity of cancer
XX therapeutics
PS
XX Example 22; Page 116-117; 173pp; English.

CC The present sequence is a partial a FRIL (Flk2/Flt3 tyrosine kinase
CC receptor-interacting lectin) protein. The specification describes a
CC composition of one or more members of FRIL family of progenitor cell
CC preservation factors. The composition is useful for alleviating or
CC reducing the hematopoietic progenitor cell-depleting activity of a
CC therapeutic treatment, including radiotherapeutic and/or
CC chemotherapeutic treatments. Administration of FRIL compositions to a
CC patient prior to treatment of the patient with a therapeutic treatment
CC having a hematopoietic progenitor cell-depleting activity alleviates or
CC reduces the hematopoietic progenitor cell-depleting activity of the
CC therapeutic treatment in the patient. FRIL family members are useful for
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
CC stem cells. The composition is administered to reduce progenitor cell
CC depleting effects of chemotherapeutics, so that the patient can receive
CC a higher dose of the chemotherapeutic and preferably recover from cancer.
CC It is also administered to patients having, or predisposed to developing
CC a condition where the patients hematopoietic progenitor cells are
CC depleted, such as severe combined immunodeficiency or aplastic anemia.
CC The isolated mesenchymal cells are useful for tissue repair.

XX
SQ Sequence 234 AA;

Query Match 98.7%; Score 1193; DB 22; Length 234;
Best Local Similarity 99.1%; Pred. No. 4.3e-102;
Matches 232; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSVSFTTFKFDSDQKLMFQGHITSSSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWE 60
Db 1 AQSVSFTTFKFDSDQKLMFQGHITSSSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWE 60

Qy 61 SSTVSTTETFTFOISTPYTSPGDLGAFFLAPYDVTIPPNSAGNLLGLFPNLRNS 120
Db 61 SSTVSTTETFTFOISTPYTSPGDLGAFFLAPYDVTIPPNSAGNLLGLFPNLRNS 120

Qy 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVADWDQNGKT 180
Db 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVADWDQNGKT 180

Qy 181 ATAHISYNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
Db 181 ATAHISYNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234

RESULT 2
AAW87973
XX AAW87973 standard; Protein; 264 AA.
XX
XX AAW87973;
XX
XX 13-APR-1999 (first entry)
XX
XX A lectin derived progenitor cell preservation factor.
XX
XX Lectin derived progenitor cell preservation factor; progenitor cell;
XX haematopoietic cell; cultured cell preservation; anticancer therapy;
XX myeloblastic therapy; sickle-cell anaemia; ablative therapy protection;
XX FLK2/FLT3 receptor.
XX
XX Dolichos lab lab.
XX
XX WO9859038-A1.
XX
XX 30-DEC-1998.
XX
XX 23-JUN-1998; 98WO-US13046.
XX
XX 24-JUN-1997; 97US-0881189.
XX

PA (IMCL-) IMCLONE SYSTEMS INC.
PA (REGC) UNIV CALIFORNIA.
XX
XX Chrispeels MJ, Colucci MG, Moore JG;
XX
XX WPI; 1999-081274/07.
XX N-PSDB; AAX03593.
XX
XX New nucleic acid encoding plant lectin that preserves progenitor
XX cells - particularly haematopoietic progenitors, useful for bone
XX marrow reconstitution after ablative therapy, and to increase DNA
XX transfer in gene therapy
XX
XX Claim 1; Page 30-31; 72pp; English.

XX The present sequence represents a lectin derived progenitor cell
XX preservation factor. The protein is used to preserve unipotent,
XX pluripotent or totipotent progenitor cells, especially haematopoietic
XX cells, and also progenitors from nerve, muscle, skin, gut, bone,
XX kidney, liver, pancreas or thymus. Specific applications are
XX preservation of cultured cells intended for administration after
XX (anticancer) myeloblastic therapy (bone marrow or whole-body irradiation
XX or chemotherapy) to reconstitute the haematopoietic system; enrichment
XX of progenitor cells (e.g. during ex vivo purging of malignant cells);
XX treatment of tissues containing haematopoietic progenitors for subsequent
XX transplant to improve haematopoietic competence; improving transfer of
XX exogenous DNA to progenitor cells (in gene therapy of various
XX haematological disorders, e.g. sickle-cell anaemia); and protection
XX against ablative therapy (to eliminate proliferating cells specifically),
XX followed by re-establishment of differentiation and proliferation of
XX preserved progenitors. The protein, when linked to magnetic beads, may
XX also be used to isolate cells that express the FLK2/FLT3 receptor.
XX
XX Sequence 264 AA;

Query Match 65.9%; Score 796.5; DB 20; Length 264;
Best Local Similarity 66.7%; Pred. No. 2.1e-65;
Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

Qy 1 AQSVSFTTFKFDSDQKLMFQGHITSSSNVQLTKLDSNGNPVSTSVGRVLYSAPLRLWE 60
Db 1 AQSLSFSTFTEDPQEDLIFQGHATSTNNVLOVTKLDSAGNPVSSSAGRVLYSAPLRLWE 60

Qy 61 SSTVSTTETFTFOISTPYTSPGDLGAFFLAPYDVTIPPNSAGNLLGLFPNLRNS 120
Db 61 DSAVLTSFDITINFEISTPYTSRIADGLAFTIAPPDSVI--SYHGGFLGLFPNANTLNS 118

Qy 121 TTSKETTIDVNAASNVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVADWDQNGKT 180
Db 119 STSENQQT-TTKAASSNVVAVEFDTPNDNIGDPYRKHGIDVNSIRSKATVADWDQNGKI 177

Qy 181 ATAHISYNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
Db 178 ATAHISYNSASKRLSVTTFYPPGKAVSLSHDVELTQVLPQWIRVGFSASTGQDK 231

RESULT 3
AAG62890
ID AAG62890 standard; Protein; 264 AA.
XX
XX AAG62890;
XX
XX 17-SEP-2001 (first entry)
XX
XX Amino acid sequence of a hyacinth bean FRIL polypeptide.
XX
XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
XX progenitor cell preservation factor; radiotherapy; chemotherapy;
XX progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
XX severe combined immunodeficiency; aplastic anemia; tissue repair.
XX
XX Dolichos lab lab.
XX

PN WO200149851-A1.
 XX 12-JUL-2001.
 XX 30-DEC-1999; 99WO-US31307.
 XX 30-DEC-1999; 99WO-US31307.
 XX (PHYL-) PHYLOGIX LLC.
 XX Colucci MG, Chrispeels MJ, Moore JG;
 XX WPI; 2001-441882/47.
 XX N-PSDB; AAH42287.
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -
 XX
 XX Example 1; Page 54-55; 173pp; English.
 PS
 CC The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of
 CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 264 AA;
 Query Match 65.9%; Score 796.5; DB 22; Length 264;
 Best Local Similarity 66.7%; Pred. No. 2.1e-65;
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;
 QY 1 AQSVSFTFTKFDSDQKDLMFQGHITSSNVLIQTLKLDNGNPVSTSVGRVLSAPLRLWE 60
 DB 1 AQSLSFSFTKFDPNQEDLI FQGHATSTNNVLQVTKLDSAGNPVSSAGRVLSAPLRLWE 60
 QY 61 SSTVVSFTFTFTQISTPYTPSPGDLGAFFLAPYDVTIPPNSAGNLLGLFNNALRNS 120
 DB 61 DSAVLTSTFTIINFESTPTYSRIADGLAFLAPDPSVI--SYHGFLGLFNNALRNS 118
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDIGDPYRKHIIGIDVNSIRSKATVADWQNGKT 180
 DB 119 STSENQT--TTKRASSNVVAVFEDTYPNDYDGNPYNIHIGIDVNSIRSKATVADWQNGKI 177
 QY 181 ATAHISYNSAKRLSVTTFTYPGGKAVLSHVDVELTQVLPQWIRVGFSASTGLEK 234
 DB 178 ATAHISYNSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDK 231
 RESULT 4
 AAG62894
 ID AAG62894 standard; Protein; 286 AA.
 XX
 AC AAG62894;
 XX
 DT 17-SEP-2001 (first entry)
 XX

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.
 XX
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;
 KW alpha-amylase inhibitor gene.
 XX
 OS Synthetic.
 OS Unidentified.
 OS Dolichos lab lab.
 XX
 PN WO200149851-A1.
 XX 12-JUL-2001.
 XX 30-DEC-1999; 99WO-US31307.
 XX 30-DEC-1999; 99WO-US31307.
 XX (PHYL-) PHYLOGIX LLC.
 XX Colucci MG, Chrispeels MJ, Moore JG;
 XX WPI; 2001-441882/47.
 XX N-PSDB; AAH42295.
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -
 XX
 XX Example 1; Page 59; 173pp; English.
 PS
 CC The present sequence represents fusion protein of alpha-amylase inhibitor
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting
 CC lectin). The specification describes a composition of one or more members
 CC of FRIL family of progenitor cell preservation factors. The composition
 CC is useful for alleviating or reducing the hematopoietic progenitor
 CC cell-depleting activity of a therapeutic treatment, including
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of
 CC FRIL compositions to a patient prior to treatment of the patient with
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting
 CC activity alleviates or reduces the hematopoietic progenitor
 CC cell-depleting activity of the therapeutic treatment in the patient.
 CC FRIL family members are useful for isolating population of progenitor
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is
 CC administered to reduce progenitor cell depleting effects of
 CC chemotherapeutics, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 286 AA;
 Query Match 65.9%; Score 796.5; DB 22; Length 286;
 Best Local Similarity 66.7%; Pred. No. 2.4e-65;
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;
 QY 1 AQSVSFTFTKFDSDQKDLMFQGHITSSNVLIQTLKLDNGNPVSTSVGRVLSAPLRLWE 60
 DB 23 AQSLSFSFTKFDPNQEDLI FQGHATSTNNVLQVTKLDSAGNPVSSAGRVLSAPLRLWE 82
 QY 61 SSTVVSFTFTFTQISTPYTPSPGDLGAFFLAPYDVTIPPNSAGNLLGLFNNALRNS 120
 DB 83 DSAVLTSTFTIINFESTPTYSRIADGLAFLAPDPSVI--SYHGFLGLFNNALRNS 140
 QY 121 TTSKETTIDVNAASNNVAVFEDTYPNDIGDPYRKHIIGIDVNSIRSKATVADWQNGKT 180
 DB 141 STSENQT--TTKRASSNVVAVFEDTYPNDYDGNPYNIHIGIDVNSIRSKATVADWQNGKI 199


```

Db 36 ALHEMFNQSKDQDLILOGDATTGTDGNLELTRVSSNGSPQSGSVGRALFYAPVHIWES 95
    : : | | : | | | | | : : : | | | : | | | | : | | | : | | |
QY 62 STVYSTETTTTFOISTPTSPPGDGLAFFAPYDTVTPPNSAGNLLGLFPNLALRNST 121
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 96 SAVVASEATFTFLIKSP-DSHPADGIAFFTSNIDSSIPSGSTGRLLGLFPDANVIRNS- 153
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
QY 122 TSKETTTIDVNAASN--NVVAVEFDTPNDNIGDPYRKHIGIDVNSRSTRKATVANDWONGK 179
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 154 ----TTIDFNAAYNADTIVAVELDTYPNTDGDPSYPHIGIDIKSVRSKKTAKWNQNGK 209
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 210 VGTAAHIINYNDKRLSAVSPNADSATVSDVDLDNVLPEWVRVGLSASTGLYK 264
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |

RESULT 7
AAV58737
ID AAV58737 standard; Protein; 290 AA.
XX
AC AAV58737;
XX
DT 25-APR-2000 (first entry)
XX
XX Jack bean concanavalin A N152S mutant.
XX
KW Concanavalin A; Con A; lectin; jack bean; insecticide;
KW transgenic plant; Brassica, insect resistance; mutant; mutein.
XX
XX Canavalia ensiformis.
XX
FH Key Location/Qualifiers
FT misc_difference 152
FT /note= "replaces wild-type Asn"
XX
XX WO200001223-A1.
XX
XX 13-JAN-2000.
XX
XX 02-JUL-1999; 99WO-SE01209.
XX
XX 03-JUL-1998; 98SE-0002425.
XX
XX (PLAN-) PLANT SCI SVERIGE AB.
XX
XX Alman I, Melander M, Vamling K;
XX
XX WPI; 2000-160693/14.
XX
XX N-PSDB; AAZ58018.
XX
XX Novel lectins used to produce transgenic Brassica plants which are
XX resistant to insects
XX
XX Example 2a; Fig 6; 5lpp; English.
XX
XX The present sequence is that of a jack bean concanavalin A
XX (Con A) mutein in which the Asn-152 residue of the native protein
XX is replaced by Ser in order to disrupt a glycosylation signal. The
XX present invention relates to: a transformed Brassica plant that is
XX resistant to certain insect pests; an expression cassette
XX containing DNA that codes for at least 1 lectin selected from Con A
XX (see AAY58736), modified Con A and pea lectin (see AAY58738); transgenic
XX plant cells containing at least 1 copy of the DNA; the present
XX lectin derived from the jack bean Con A gene; a method of imparting
XX resistance to insects selected from blossom beetles (pollen
XX beetles) of the genus Meligethes, flea beetles of the genus
XX Phyllotreta, and root flies of the genus Delia; and a method for
XX protecting a plant against infestation by insects of these genera.
XX
XX Sequence 290 AA;

```

Query Match 52.8%; Score 638.5; DB 21; Length 290;
 Best Local Similarity 53.6%; Pred. No. 9.4e-51;

```

Matches 126; Conservative 36; Mismatches 64; Indels 9; Gaps 4;
QY 3 SVSFTTKFSDQKDLMFQGH-TISSNNVIOLTKLDSNCPVSTVGRVLVSAPRLWES 61
    : : | | : | | | | | : : : | | | : | | | | : | | | : | | |
Db 36 ALHEMFNQSKDQDLILOGDATTGTDGNLELTRVSSNGSPQSGSVGRALFYAPVHIWES 95
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
QY 62 STVYSTETTTTFOISTPTSPPGDGLAFFAPYDTVTPPNSAGNLLGLFPNLALRNST 121
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 96 SAVVASEATFTFLIKSP-DSHPADGIAFFTSNIDSSIPSGSTGRLLGLFPDANVIRNS- 153
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
QY 122 TSKETTTIDVNAASN--NVVAVEFDTPNDNIGDPYRKHIGIDVNSRSTRKATVANDWONGK 179
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 154 ----TTIDFNAAYNADTIVAVELDTYPNTDGDPSYPHIGIDIKSVRSKKTAKWNQNGK 209
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
QY 180 TATAHISYNSASKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTGLEK 234
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |
Db 210 VGTAAHIINYNDKRLSAVSPNADSATVSDVDLDNVLPEWVRVGLSASTGLYK 264
    | | | | | | | | : | | | | | : | | : | | | | | : | | | |

RESULT 8
AAV74765
ID AAR74765 standard; protein; 237 AA.
XX
AC AAR74765;
XX
DT 13-OCT-1995 (first entry)
XX
DE Legume concanavalin A.
XX
KW Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;
KW amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;
KW female hamster protein; 3-D structure; jellyroll topology;
KW picornavirus; coat protein; pentameric structure; concanavaline A.
XX
XX Legume sp..
XX
XX WO9505394-A.
XX
XX 23-FEB-1995.
XX
XX 17-AUG-1994; 94WO-GB01802.
XX
XX 17-AUG-1993; 93GB-0017120.
XX
XX (BIRK-) BIRKBECK COLLEGE.
XX
XX (ROYA-) ROYAL POSTGRAD MED SCHOOL.
XX
XX Blundell TL, Pepys MB;
XX
XX WPI; 1995-098720/13.
XX
XX New cpds. Inhibiting binding of serum amyloid P to amyloid
XX fibrils - produced by computer assisted molecular design, useful
XX for preventing, treating or diagnosing amyloidosis, e.g.
XX Alzheimer's disease
XX
XX Claim 7; Fig 4; 72pp; English.
XX
XX The sequences given in AAR74763-70 represent various pentraxins and
XX legume lectins. All these proteins show structural homology and may
XX be used in the molecular design of a molecule for the inhibition of
XX serum amyloid P (SAP) binding to amyloid fibrils. The similarities
XX in the amino acid sequences of SAP, human and Limulus C-reactive
XX protein (CRP) and female hamster protein suggests that they may have
XX similar 3-D structures. The jellyroll topology of the pentraxins is
XX reminiscent of the picornavirus coat proteins which also have
XX pentameric structures. However, pentaxins resemble more closely
XX legume lectins such as concanavaline A and pea lectin. Alignment
XX of sequences on the basis of topologically equivalent features of the
XX three dimensional structures shows that helices occupy different
XX positions in the pentraxins and legume lectins and that the amino acid
XX sequences of the two families have identities of only approx. 11%. The
XX two main helices in SAP occur before and after strand L, whereas the

```

[illegible]

AC	XX	AA100611;
XX	XX	
DT	XX	23-JUN-1999 (first entry)
DE	XX	
XX	XX	R. pseudoacacia lectin RPAI polypeptide subunit a.
KW	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
KW	XX	cancer; metabolic disorder; mucositis; cytotoxic.
XX	XX	
OS	XX	Robinia pseudoacacia.
XX	XX	
PN	XX	WO9911278-A1.
PD	XX	11-MAR-1999.
XX	XX	
PF	XX	28-AUG-1998; 98WO-GB02612.
XX	XX	
PR	XX	29-AUG-1997; 97GB-0018413.
XX	XX	
PA	XX	(ALIZ-) ALIZYME THERAPEUTICS LTD.
PI	XX	Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
XX	XX	
DR	XX	WPI; 1999-228935/19.
XX	XX	
PT	XX	Use of Robinia pseudoacacia lectin in medicine ;
XX	XX	
PS	XX	Disclosure; Fig 1; 37pp; English.
XX	XX	

CC The invention relates to Robinia pseudoacacia lectins which can be used
 CC in medicine. The lectin is used for the control of mucosal cell
 CC proliferation, for the reduction and/or treatment of damage caused by
 CC cell damaging agents, especially in the treatment of cancer, and/or for
 CC the reduction and/or treatment of metabolic disorders. It is especially
 CC useful for the treatment of mucositis in mammalian cells and/or tissues
 CC particularly human cells and/or tissues (especially mucous cells
 CC including mucous membrane). Use of the lectin is effective and does not
 CC damage the gut as is the case with cytotoxic drugs and radiation. The
 CC present sequence represents a polypeptide subunit of *R. pseudoacacia*
 CC lectin RbAl.
 xx
 sq Sequence 285 AA;

[illegible]

RESULT 11	
AAAY06814	
ID	AAAY06814 standard; protein; 285 AA.
XX	
AC	AAAY06814;
XX	
XX	
DT	23-JUN-1999 (first entry)
XX	
DE	Amino acid sequence of lectin RPAI from R. pseudoacacia seed.
XX	
KW	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
KW	cancer; metabolic disorder; mucositis; cytotoxic.
XX	
OS	Robinia pseudoacacia.
XX	
PN	WO9911278-A1.
XX	
PD	11-MAR-1999.
XX	
PF	28-AUG-1998; 98WO-GB02612.
XX	
PR	29-AUG-1997; 97GB-0018413.
XX	
PA	(ALIZ-) ALIZYME THERAPEUTICS LTD.
XX	
PI	Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
XX	
DR	WPI: 1999-228935/19.

Use of Robinia pseudoacacia lectin in medicine
Disclosure: Fig 2; 37pp; English.

The invention relates to Robinia pseudoacacia lectins which can be used in medicine. The lectin is used for the control of mucosal cell proliferation, for the reduction and/or treatment of damage caused by cell damaging agents, especially in the treatment of cancer, and/or for cell reduction and/or treatment of metabolic disorders. It is especially useful for the treatment of mucositis in mammalian cells and/or tissues, particularly human cells and/or tissues (especially mucous cells including mucous membrane). Use of the lectin is effective and does not damage the gut as is the case with cytotoxic drugs and radiation. The present sequence represents a R. pseudoacacia lectin RPA1. This lectin is composed of two subunits- polypeptide a and polypeptide b (AA056811-12).

Sequence 285 AA:

Query Match 36.8%; Score 445; DB 20; Length 285;
Best Local Similarity 41.9%; Pred. NO. 6.8e-33;
Matches 98; Conservative 49; Mismatches 65; Indels 22; Gaps 8;

```

Qy 3 SVSFTTKFDSOKDLMFQGH-T-ISSSNVIOITKLDSNGNPVSTSVGRVLVYSAPLRWES 61
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 34 SLUSFPPKAPNPXYLIQFDALVTSGVLGTNV-VNCGVPPRRSIGRALLYAAPQTQDN 92
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 62 ST-VVSTFETTFQISPTPTSPGDDGLAFLAPYDVTVPNSAGNLLGLFPNLNLRNS 120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 93 TTGNVASEVTSFIIQAPNPATTADGLAFLAPYDVT--QPGDLGMLGTIF-----KDG 144
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 TTSKETIDVNAASNWVAFEDTYPNDNIGDPYRKHHIGIDVNSIRSKATVAMDWONGKT 180
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 145 SYNK-----SNQIVAVEFDFDSNIHF-DPGRHMGINVSIVSVKTPVNPWTNGEV 194
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 ATAHTSYNSASKSLSVTTFYPGGKAVLSLSD-VELTQVLPQWIRVGFSGASTGLE 233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 ANVFISYEASTSLNASLVPSLETSTFIHALVDVKDVLPEWVRFGFSATTGID 248
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
AAY06815
ID AAY06815 standard; protein; 285 AA.

```

RESULT 12	
XX AAY06815	standard; protein; 285 AA.
XX ID AAY06815	
XX AC	AAY06815;
XX XX	
XX DT	23-JUN-1999 (first entry)
XX XX	
XX DE	Amino acid sequence of lectin RPBHII from R. pseudoacacia seed.
XX XX	
XX RW	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
XX KW	cancer; metabolic disorder; mucositis; cytotoxic.
XX XX	
XX OS	Robinia pseudoacacia.
XX XX	
XX PN	WO9911278-A1.
XX PD	
XX PF	11-MAR-1999.
XX PR	
XX XX	28-AUG-1998; 98WO-GS02612.
XX PA	
XX PI	29-AUG-1997; 97GB-0018413.
XX DR	(ALIZ-) ALIZYME THERAPEUTICS LTD.
XX PT	Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
XX PS	WPI; 1999-228935/19.
	Use of Robinia pseudoacacia lectin in medicine
	Disclosure: Fig 2; 37pp; English.

The invention relates to Robinia pseudoacacia lectins which can be used
 in medicine. The lectin is used for the control of mucosal cell
 proliferation, for the reduction and/or treatment of damage caused by
 cell damaging agents, especially in the treatment of cancer, and/or for
 the reduction and/or treatment of metabolic disorders. It is especially
 useful for the treatment of mucositis in mammalian cells and/or tissues,
 particularly human cells and/or tissues (especially mucous cells
 including mucous membrane). Use of the lectin is effective and does not
 damage the gut as is the case with cytotoxic drugs and radiation. The
 present sequence represents a R. pseudoacacia lectin RPB11. This lectin
 is composed of the subunit- polypeptide c (AAY06813).
 Sequence 285 AA:
 XX
 SO

Query Match	36.7%	Score 444:	DB 20:	Length 285:
Best Local Similarity	42.7%	Pred. NO. 8.5e-33:		
Matches 100: Conservative	46:	Mismatches 66:	Indels 22:	Gaps 8:

Qy	3	SYSFTTKFDSOKDLMFQOHT - ISSSNVIQTLKLDNSGNPNVSTSYGVRVLYSAPLRWES	61
		: : : : : : : : : : : : : : : : : : :	
Db	34	SLSFSPKFAPNQPYLIFQDALVTSTGVLQTNV - VNGVPSRKSIGRALYAAPFIQWDS	92
Qy	62	ST - VSTFEETFTTQISTPTSPGPDGIAFFLAPYDVTVIPPNSAGNLLGLFPNLNALRNS	120

Db 93 TTNVASFVTSFIQAPNPATAGLAFFLAPVDI--OPLDGGMLGIF-----KNG 144
 QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYR--KHIGIDVNSIRSKATVAMWQNGKT 180
 Db 145 YFNK-----SNQIVAVEFDTFESNRH-WDPTGRHMGINVNSIVSVKTPWPNWANCEV 194
 QY 181 ATAHISYNSASKRLSVTTTPYGGKAVSLSHD--VELTQVLVQWIRVGFSASTGLE 233
 Db 195 ANVFISYEASTKSLTASLVPSLETSLIIHAIIVDKVDVLPWVRFVGFSAATTGID 248

RESULT 13

RAY58738
 ID AAY58738 standard; Protein; 275 AA.

XX AC AAY58738;
 XX DT 25-APR-2000 (first entry)

XX DE Pea lectin.

XX KW Lectin; pea; insecticide; transgenic plant; Brassica;
 XX KW insect resistance.
 XX OS Pisum sativum.

XX FN WO200001223-A1.

XX PD 13-JAN-2000.

XX PF 02-JUL-1999; 99WO-SE01209.

XX PR 03-JUL-1998; 98SE-0002425.

XX PA (PLAN-) PLANT SCI SVERIGE AB.

XX PI Ahman I, Melander M, Vamling K;

XX DR WPI; 2000-160693/14.

XX DR N-PSDB; AAZ58019.

XX Novel lectins used to produce transgenic Brassica plants which are
 resistant to insects
 Example 2b; Fig 9; 51pp; English.

XX The present sequence is that of pea cv. Lincoln lectin. The
 invention relates to: a transformed Brassica plant that is
 resistant to certain insect pests; an expression cassette
 containing DNA that codes for at least 1 lectin selected from Con A
 (see AAY58736), modified Con A (see AAY58737) and pea lectin;
 transgenic plant cells containing at least 1 copy of the DNA; a new
 lectin derived from the jack bean Con A gene; a method of imparting
 resistance to insects selected from blossom beetles (pollen
 beetles) of the genus *Meligethes*, flea beetles of the genus
Phyllotreta, and root flies of the genus *Delia*; and a method for
 protecting a plant against infestation by insects of these genera.

XX SQ Sequence 275 AA;

Query Match 34.0%; Score 411.5; DB 21; Length 275;
 Best Local Similarity 40.2%; Pred. No. 8.1e-30;
 Matches 9; Conservative 39; Mismatches 72; Indels 33; Gaps 8;

QY 2 QSVSFTTKFSDOKLDFQGHITSSSNVQLKDSNGNPVSTSVGRVLYSAPLRLWES 61

Db 32 ETTSELTIKFSPDQNLIFQGDGYTTREKLTITK-----AVKNTVGRALYSSPHIHDR 85

QY 62 ST-VVSTFTFTTFTTQISTPYTSPGDDGLAFFLAPYDVTVPNPSAGNLLGLFPNLNLR 120

Db 86 ETGNVANEVTSFTVINAPNSYNVADGTFETFTIAPVDI--KPQTGGGYLGVE-----NS 136

QY 121 TTSKETTIDVNAASNNVAVVEFDTPNDNIGDPYR--KHIGIDVNSIRSKATVAMWQNG 178
 Db 137 AEYDKTT-----QTVAVEFDTFYN-AAWDPNRRDHIGIDVNSIKSVNTKSWKLQNG 187
 QY 179 KTAHAHSYNSASKRLSVTTTPYGG-----KAVSLSHDVELTQVLVQWIRVGFSASTGL 232
 Db 188 EAVNVIAFNAATNVLTVSLTPNPSLEENVTSTYTLSDVSLKDVVPEWVRFVGFSAATTGA 247
 QY 233 E 233
 Db 248 E 248

RESULT 14

AAR45911
 ID AAR45911 standard; protein; 242 AA.

XX AC AAR45911;

XX DT 18-JUL-1994 (first entry)

XX DE Bauhinia purpurea lectin larvicidal deriv.

XX KW Non-lysine; insect larvae; sorghum; wheat; oat; rye; rice; European;
 XX KW corn borer.
 XX OS Bauhinia purpurea.

XX PN WO9402514-A.

XX PD 03-FEB-1994.

XX PF 22-JUL-1993; 93WO-US06946.

XX PR 24-JUL-1992; 92US-0921179.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Balasubramaniam NK, Rao AG;

XX DR WPI; 1994-048798/06.

XX Bauhinia purpurea lectin larvicidal non-lysine derivs. - and
 PT corresp. DNA used to transform e.g. maize plants, to protect
 PT plants or stored prods. from insect larvae, e.g. European corn
 PT borer
 XX Claim 1; Page 17; 19pp; English.

XX The sequence is that of a Bauhinia purpurea lectin which has had
 CC lysine residues replaced by other amino acids which either preserve
 CC the positive charge at the position of the substitution or provide a
 CC neutral residue. The protein derivs. are effective larvicides
 CC against insect pupae, such as the European corn borer. The protein
 CC can be administered enterally to the larvae in their diet and can be
 CC used to protect sorghum, wheat, oats, rye, etc.
 CC See also AAR45912.

XX SQ Sequence 242 AA;

Query Match 31.3%; Score 379; DB 15; Length 242;
 Best Local Similarity 37.0%; Pred. No. 6.8e-27;
 Matches 88; Conservative 47; Mismatches 77; Indels 26; Gaps 7;

QY 6 FTFTKFDSDOK-----DLMFOGHITSSSNVQLKDSNGNPVSTSVGRVLYSAPLRLWES 61

Db 8 FTFPNFWSTFQENGTELIIFLGATYTPGALRLRIGEDGIPKLSNAGQASYSRPFVLWDS 67

QY 62 STVVSFTFTTFTF---QISTPYTSPGDDGLAFFLAPYDVTVPNPSAGNLLGLFPNLNLR 118

Db 68 TGHVASFYTSFIVRSIDVPHIT--ADGFAFLAPDVSSV--KDYGGCLGLF-----R 117

QY 119 NSITSKETTIDVNAASNNVAVVEFDTPNDNIGDPYRKKHIGIDVNSIRSKATVAMWQNG 178

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 ; Search time 8.17978 Seconds
(without alignments)
1079.114 Million cell updates/sec

Title: US-09-476-485A-8

Perfect score: 1209

Sequence: 1 AQSVSFTFTKFDSDQKDLMF.....TQVLPOWIRGVFSASTGLEK 234

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	20.6	669	10	US-09-862-027-27
2	101	8.4	2383	10	US-09-912-020-302
3	97	8.0	2344	10	US-09-815-242-12713
4	94.5	7.8	345	10	US-09-813-820-7
5	90.5	7.5	1169	10	US-09-801-368-106
6	84.5	7.0	448	10	US-09-871-212-8
7	83.5	6.9	888	10	US-09-826-752-6
8	83.5	6.9	1723	10	US-09-841-132-394
9	83.5	6.9	1723	10	US-09-841-132-395
10	83.5	6.9	5795	10	US-09-815-242-12610
11	83	6.9	283	10	US-09-925-299-984
12	83	6.9	595	10	US-09-881-752A-166
13	81.5	6.7	599	9	US-09-771-382-6
14	81.5	6.7	599	10	US-09-787-862-15
15	80.5	6.7	1367	10	US-09-801-368-108
16	80.5	6.7	2092	9	US-10-147-026-12
17	79.5	6.6	507	9	US-09-738-626-5458
18	79.5	6.6	922	10	US-09-886-468-19
19	79	6.5	906	10	US-09-905-983-46

20	79	6.5	906	10	US-09-746-491-46	Sequence 46, Appl
21	79	6.5	2026	10	US-09-801-368-86	Sequence 86, Appl
22	79	6.5	2224	9	US-10-115-563-14	Sequence 14, Appl
23	78.5	6.5	1848	9	US-09-839-996-6	Sequence 6, Appl
24	78	6.5	413	10	US-09-746-491-43	Sequence 43, Appl
25	77	6.4	422	9	US-09-738-626-5947	Sequence 5947, Ap
26	77	6.4	969	9	US-09-981-353-122	Sequence 122, App
27	77	6.4	977	10	US-09-925-297-797	Sequence 797, App
28	76.5	6.3	632	10	US-09-853-533A-8	Sequence 8, Appl
29	76.5	6.3	1447	10	US-09-797-097-2	Sequence 2, Appl
30	76	6.3	579	10	US-09-870-203A-6	Sequence 6, Appl
31	76	6.3	579	10	US-09-870-203A-8	Sequence 8, Appl
32	76	6.3	579	10	US-09-870-203A-18	Sequence 18, Appl
33	76	6.3	579	10	US-09-870-203A-10	Sequence 20, Appl
34	76	6.3	580	10	US-09-870-203A-12	Sequence 12, Appl
35	76	6.3	581	10	US-09-758-008-5	Sequence 5, Appl
36	76	6.3	581	10	US-09-870-203A-2	Sequence 2, Appl
37	76	6.3	581	10	US-09-870-203A-4	Sequence 4, Appl
38	76	6.3	581	10	US-09-870-203A-14	Sequence 14, Appl
39	76	6.3	581	10	US-09-870-203A-16	Sequence 16, Appl
40	76	6.3	600	10	US-09-871-212-5	Sequence 5, Appl
41	76	6.3	600	10	US-10-074-279-6	Sequence 6, Appl
42	75.5	6.2	451	9	US-10-074-279-6	Sequence 22, Appl
43	75.5	6.2	594	10	US-09-815-108-22	Sequence 4, Appl
44	75.5	6.2	660	9	US-09-872-462-4	Sequence 218, Appl
45	75.5	6.2	704	10	US-09-801-368-218	

ALIGNMENTS

RESULT 1
US-09-862-027-27
; Sequence 27, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: NO. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-862-027-27

Query Match		20.6%	Score 249;	DB 10;	Length 669;
Best Local Similarity		32.8%	Pred. No. 7.3e-15;		
Matches		79;	Conservative 41;	Mismatches 81;	Indels 40; Gaps 10;
Qy	13	SDQKDLMFQHTTSSNNVQLTKLDSNG-----NPVSTVGRLVYSAPLRLWES-STVVS	66		
Db	23	SEEGFNGVLYDMSG-IAIT--NSKGLMKLTNSEFSYGHVFNPSVFRNPNMGTVS	79		
Qy	67	TFETTFQISTPYTSPPGDGLAFFLAPYDVTPPNAGNLLGLFNLNLRNNTTTSKET	126		
Db	80	SFSTTFEFAIVSNVNALDGHGLAFVISP-TKGLPYSSSSQYLCLF-----NLTNNGD-	130		
Qy	127	TIDVNASNNVAVVEFDYPNDNIGDPYRKHGIDVNSIRS-KAIVAWDQ-----	176		
Db	131	-----PSNHVAVVEFDYFQNOEFDMDNNHVGIDINLSLSEKASTAGYEDDGTGFKNI	184		
Qy	177	---NCKTATAHISYNSAKRLSVT---TFYGGKAVSLSHDVELTQVLPQWLVRGVFSAST	230		
Db	185	RLINQKPIQAWIEWDSSRQLNVTIHPILPKPIPLLSLTLDKLSPLFDSMVIVGFTSAT	244		
Qy	231	G 231			

Db 245 G 245

RESULT 2

US-09-912-020-302

; Sequence 302, Application US/09912020

; Patent No. US20020045592A1

; GENERAL INFORMATION:

; APPLICANT: Zyskind, Judith

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Trawick, John

; APPLICANT: Forsyth, R. Allyn

; APPLICANT: Froelich, Jamie M.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN

; FILE REFERENCE: ELITRA.001DV1

; CURRENT APPLICATION NUMBER: US/09/912,020

; CURRENT FILING DATE: 2001-07-23

; PRIOR APPLICATION NUMBER: 09/492,709

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: 60/117,405

; PRIOR FILING DATE: 1999-01-27

; NUMBER OF SEQ ID NOS: 485

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 302

; LENGTH: 2383

; TYPE: PRT

; ORGANISM: E. Coli

; US-09-912-020-302

Query Match

Best Local Similarity 8.4%; Score 101; DB 10; Length 2383;

Matches 64; Conservative 37; Mismatches 101; Indels 96; Gaps 11;

QY 13 SDOKLDFOGHTI-----SSNVQLTKLDSNGNPVS-----44

Db 1257 SDNKTVEHFGDFAAKIIEIAPVPSIIAGTPQNSGSGSVITATVVDNNGSPVKGVTVNF 1316

QY 45 -----TSVGR-----VLYSAPLRLEWSTVSTFTFTTFTQISTPTVTPPGD 86

Db 1317 SNAETAEMTNGQAVNEOGKATVTVNTRSSIESCARPTVASELNGSSTLSTSNVN 1376

QY 87 G-----LAFFLAPYDVTIPPSAGNLLGLFPNL-NALRNSSTSKETIDVNA-----S 134

Db 1377 ADASTAHLTLQALPDTV---SAGETTSLYIEVKDNYGNGVPOQEVTLSPSPSEGVTPS 1432

QY 135 NNVAVEFTDYNIDGDPYRKHIGIDVNSIRSKAIVAMDQWQKTATAHISY-----187

Db 1433 NNAI-----YTNHNDGNFYASFTATKAGVYQLTATL-----ENGDSMQQVTVYVPVANA 1482

QY 188 -----NSAKRLSVTTFYPGKAVSLSHDVELTQVLPQWIRVGFSASTG 231

Db 1483 EITLAASKDPVADNNDLTLTATVADTEGNAIA---NTEVTTLPEDEVKANTLSDG 1537

RESULT 3

US-09-815-242-12713

; Sequence 12713, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari L.

; APPLICANT: Zyskind, Judith W.

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John D.

; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.

; APPLICANT: Xu, H. Howard

; TITLE OF INVENTION: Identification of Essential Genes in

; TITLE OF INVENTION: Prokaryotes

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 12713

; LENGTH: 2344

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

; US-09-815-242-12713

Query Match 8.0%; Score 97; DB 10; Length 2344;

Best Local Similarity 24.2%; Pred. No. 3.2;

Matches 58; Conservative 33; Mismatches 85; Indels 64; Gaps 12;

QY 2 QSVFTFTKFDSDQKDLMFQGHTI---SSNVQLTKLDSNGNPVSTSVGRV-----50

Db 565 QSVYYFTI--DVKAPTIVGNGQTEVGKTMNPVILTTDNGTGTVNTVGLPSGLSYDS 622

QY 51 ----LYSAPLRLEWSS-TVSTFE-----TFTTFOI---STPYTSPGDDGLAFLAPYD 96

Db 623 ATNSIIGTPTKIGQSTVTVSTDOANNKSTTTFTTNVVDTTAPTPTVIGDKSSEVFSPIS 682

QY 97 TVIPNSAGNLLGLFPNLNALRNSITTSKETIDVNAASNNVV-----AVEFDTPNDNIGD 152

Db 683 PI-----NATQDNGS-----NAVNTVTVGLPSGLFDSTNTTIST 719

QY 153 PYRKHIGIDVNSIRSKATVAMDQWQKTATA---HISYNSASKRLSVTTFYPGKAVSL 209

Db 720 P--INIGTSTITI-----VSTDASGNKTTTTFKVEVTRNMSDSVSTSGTQOSQSVST 772

RESULT 4

US-09-813-820-7

; Sequence 7, Application US/09813820

; Patent No. US20020102262A1

; GENERAL INFORMATION:

; APPLICANT: Hook, Magnus

; APPLICANT: Patti, Joseph M.

; APPLICANT: House-Pompeo, Karen

; APPLICANT: Sthanam, Narayana

; APPLICANT: Symersky, Jindrich

; TITLE OF INVENTION: COLLAGEN BINDING PROTEIN COMPOSITIONS

; AND METHODS OF USE

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: Texas

; COUNTRY: U.S.

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/813,820

```
; FILING DATE: 22-Mar-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/856,253
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitchell, Barbara S.
; REGISTRATION NUMBER: 33,928
; REFERENCE/DOCKET NUMBER: TANK:193
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 345 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-813-820-7

Query Match          7.8%; Score 94.5; DB 10; Length 345;
Best Local Similarity 24.18; Pred. No. 0.41;
Matches 59; Conservative 32; Mismatches 73; Indels 81; Gaps 16;

QY 3 SVSFTFTKFDSDQKDLMPQGHITSSNNVIQLTKLDSNGN-PVSTSVGRVL----- 51
Db 107 NVIYFTFDVYNTKDDVK---ATLMPAIDPENVKKTGNVLATIGTGSTANKTVLVDYE 163
QY 52 -YSAPLRWESSTVVSFTFTFQISTPTSPYSPGGLAFLAPYDVTIPPNSAGNLLGL 110
Db 164 KYGFYNSLTKGTIDQIDKNTNTYR-QTIYVNPSC-----DNVIAPVLTGNLK-- 210
QY 111 FPNL--NAL--RNSTSKETIDVNA-----SNNVAVE 141
Db 211 -PNTDSNALIDQNTSIKYKVD-NARDLSFYFVNPENFEDVTNSVITPPNHOYKVE 268
QY 142 FDTYPNDNIGDPYRKHTG--IDVNS-----KATVAV---DWQNGKTATAHIS 186
Db 269 FNT-PDQITTPYIVVGVNHIDPNKSGDLARSTLYGYNSNIWRMSWDN-----EVA 321
QY 187 YNSAS 191
Db 322 FNGCS 326

RESULT 5
US-09-801-368-106
; Sequence 106, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 106
; LENGTH: 1169
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-106

Query Match          7.5%; Score 90.5; DB 10; Length 1169;
Best Local Similarity 21.3%; Pred. No. 4.9;
Matches 55; Conservative 32; Mismatches 112; Indels 59; Gaps 9;

QY 2 QSVSFTFTKFDSDQKDLMPQGHITSSNNV-IQLTKLDSNGNPNVSTSVGRVLYSAPLRW 60
Db 498 ESTSYVTPYVSSSTAAANYTSSFSSEVCTETES-----TSTSTPYVTSS---WS 549
QY 61 SSTVSTFTFTFQISTPTSPYSPGGLAFLAPYDVTIPPNSAGNLLGLFPNLALRNS 120
Db 550 SSEVCTETETESTSYVTPYVS-----SSTAAANYTSSFSSEVCTE 592
QY 121 TTSKEIT-----IDVNAASNVA--VEFDTPNDNIGDPYRKHHIGIDVNS 164
Db 593 CTETESTSTPYATSTGATSTASTNTMTSLVQDTTVSFSLSTSVSEHTNAPTSS 652
QY 165 IRSKATVAMDQNGKTATAHISYNSASKRLSVTTFYPGGKAVSLSHDYEL----- 214
Db 653 VESNASTFISNKGSKV-----SYVTSSIH-SITPMYPSNOTVTSSSVVSTPITSESS 707
QY 215 --TQVLPQWIRVGFSAST 230
Db 708 ASVTILPSTITSEFKPST 725

RESULT 6
US-09-871-212-8
; Sequence 8, Application US/09871212
; Patent No. US20020034519A1
; GENERAL INFORMATION:
; APPLICANT: Tikoo, Suresh
; APPLICANT: Babluk, Lorne
; APPLICANT: Zhang, Linong
; APPLICANT: Wu, Qiaohua
; TITLE OF INVENTION: MODIFIED BOVINE ADENOVIRUS HAVING
; FILE REFERENCE: 293102003000
; CURRENT APPLICATION NUMBER: US/09/871,212
; PRIOR APPLICATION NUMBER: 60/208,678
; PRIOR FILING DATE: 2000-05-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Porcine adenovirus 3
US-09-871-212-8

Query Match          7.0%; Score 84.5; DB 10; Length 448;
Best Local Similarity 25.2%; Pred. No. 4.8;
Matches 36; Conservative 23; Mismatches 63; Indels 21; Gaps 6;

QY 18 LMFQGHITSSNNVIQLTKLDSNGNPNVSTSVGRVLYSAPLRWESSTVSTFTFTFQIS 77
Db 294 LVLGSAFKASS--IDLTSMTKKVNFIFDGAGRL-----QSD---SIYKGFGRFSN 339
QY 78 TPTTTPPGDGLA-FFLAPYDVTIPPNSAGNLLGLFPNLALRNSTTSKETTIDVNAASN 136
Db 340 DSVIEPTAAGLSPAWLMPSTFIYPRNTSGSLTSFVYN-----QTVVHVHDIKVTLSN 394
QY 137 VVAVEFDTPNDNIGDPYRKHHIG 159
Db 395 GYSLEEN-FQNMSFSAPFSTSYG 416
```

RESULT 7
 US-09-826-752-6
 ; Sequence 6, Application US/09826752
 ; Patent No. US20010026930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Guarente, Leonard P.
 ; APPLICANT: Austriaco Jr., Nicanor
 ; APPLICANT: Claus, Francesca
 ; APPLICANT: Cole, Jamesca
 ; APPLICANT: Kennedy, Brian
 ; TITLE OF INVENTION: GENES DETERMINING CELLULAR SENESCENCE IN YEAST
 ; FILE REFERENCE: 0050.1491-005
 ; CURRENT APPLICATION NUMBER: US/09/826,752
 ; PRIOR FILING DATE: 2001-04-05
 ; PRIOR APPLICATION NUMBER: US 08/396,001
 ; PRIOR FILING DATE: 1995-02-28
 ; PRIOR APPLICATION NUMBER: PCT/US94/09351
 ; PRIOR FILING DATE: 1994-08-15
 ; PRIOR APPLICATION NUMBER: US 08/107,408
 ; PRIOR FILING DATE: 1993-08-16
 ; PRIOR APPLICATION NUMBER: US 09/323,433
 ; PRIOR FILING DATE: 1999-06-01
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 888
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-826-752-6

Query Match 6.9%; Score 83.5; DB 10; Length 888;
 Best Local Similarity 24.8%; Pred. No. 15;
 Matches 53; Conservative 33; Mismatches 87; Indels 41; Gaps 11;
 QY 11 FDSQKDLMFQGHITSSNVQLTKLDSNGNPV-----STSGRVLYSAPRLWESSTV-- 64
 Db 118 FRND--SLPQ---MLSSGAATGQGNLPLNDNSMKVLPASADP--LWTHSNVPG 170
 QY 65 ---VSTFTFTTQISTPTSPGCDGLA-----FFLAPYDVIPNSAGNLLGLFPN 113
 Db 171 SASVAIEETATLQESLPKSGRESNKNKASFRQTFHALSPTDLI---NAANNVTLKDF 227
 QY 114 LNALNSTSKETIDV--NAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVA 172
 Db 228 QSDQNFSAKKEKSVGANNTAKTFTQSIISDNTPSSTSFIP-----PTNSVSEKLS-- 278
 QY 173 WDQNGKTATAHISYNSAKRLSVTTF---YPGG 203
 Db 279 -DFKIETSKEDLINKTAPAKKESPTTYGAAYPYG 311

RESULT 8
 US-09-841-132-394
 ; Sequence 394, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 394
 ; LENGTH: 1723
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-841-132-394

Query Match 6.9%; Score 83.5; DB 10; Length 1723;
 Best Local Similarity 22.0%; Pred. No. 36;
 Matches 52; Conservative 26; Mismatches 85; Indels 73; Gaps 9;
 QY 1 AQSVSFTFTKFDSDOKDLMFQGHITSSNVQLTKLDSNG-----NPVSTSGRVLYSAP 55
 Db 1162 AGNVKFTAIEASA-----GKAISFYDAVNVTRETNQAELKLNKATSTGTLFSGE 1213
 QY 56 LRLWESSTVVS---TFE-----TFTFTQISTPTSPGCD-----GL 88
 Db 1214 LH--ENKSYIPQKVTFAGHNLILGKNAELSVVSTQSPGTTITMGPGSVLSNHSKEAGGI 1271
 QY 89 AFLAPYD-----TVIPP-----NSAGNLLGLFPNNAALRNST 121
 Db 1272 AINNVIIDFSEIVPTKDNATVAPPTKLKVRNADSKDKIDITGTVTLDPNGNLYQNSY 1331
 QY 122 TSKETIDV---NAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVAW 173
 Db 1332 LGEDRDITLFDNDSASGAVTATNVTLQNLGAKKGVLGTWNLDPNSSGSKILKW 1387
 RESULT 9
 US-09-841-132-395
 ; Sequence 395, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841,132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 395
 ; LENGTH: 1723
 ; TYPE: PRT
 ; ORGANISM: Chlamydia pneumoniae
 US-09-841-132-395

Query Match 6.9%; Score 83.5; DB 10; Length 1723;
 Best Local Similarity 22.0%; Pred. No. 36;
 Matches 52; Conservative 26; Mismatches 85; Indels 73; Gaps 9;
 QY 1 AQSVSFTFTKFDSDOKDLMFQGHITSSNVQLTKLDSNG-----NPVSTSGRVLYSAP 55
 Db 1162 AGNVKFTAIEASA-----GKAISFYDAVNVTRETNQAELKLNKATSTGTLFSGE 1213
 QY 56 LRLWESSTVVS---TFE-----TFTFTQISTPTSPGCD-----GL 88
 Db 1214 LH--ENKSYIPQKVTFAGHNLILGKNAELSVVSTQSPGTTITMGPGSVLSNHSKEAGGI 1271
 QY 89 AFLAPYD-----TVIPP-----NSAGNLLGLFPNNAALRNST 121
 Db 1272 AINNVIIDFSEIVPTKDNATVAPPTKLKVRNADSKDKIDITGTVTLDPNGNLYQNSY 1331
 QY 122 TSKETIDV---NAASNNVAVVEFDTPNDNIGDPYRKHIGIDVNSIRSKATVAW 173
 Db 1332 LGEDRDITLFDNDSASGAVTATNVTLQNLGAKKGVLGTWNLDPNSSGSKILKW 1387
 RESULT 10
 US-09-815-242-12610
 ; Sequence 12610, Application US/09815242
 ; Patent No. US20020061569A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John D.

Best Local Similarity 18.8%; Pred. No. 13;
Matches 40; Conservative 43; Mismatches 75; Indels 55; Gaps 10;

QY 12 DSDQKDLMFOGHTISSNNVIQLTKLDSNGNPVSTSVGRV-----LYSAPRL-- 58
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 53 DEDEEEL-EPVRSALVLQP-MIDKEGNGENESTGNIGWSIYYDNHNTLHGATVTLKA 109

QY 59 -----WESSTVVVFETFTFQIS---TPYSPGDGLAFFLPDYTVIPNSAGN 106
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 110 GDNLIKONTKNENTNDSSFTSLKKDLTDLTSVETEKLSF-----CANGN 158

QY 107 LLGLFPNLNAL----RNSTTSKETTTIDVNAASNVAVEFDT-----YPNDNIGDPYRKH 157
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 159 KVNITSDIKGLNFAKETAGTNGDTTVHLNGIGSTLTDLTLTGATTNVTDNDVDDKKR 218

QY 158 IGIDVNSIRSKATVAWDWN---GKTATAHSY 187
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 219 AA-----SVKDVLNAGWNIIKGVKPGTTASDNDVF 247

RESULT 15
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

Query Match 6.7%; Score 80.5; DB 10; Length 1367;
Best Local Similarity 22.6%; Pred. No. 50;
Matches 47; Conservative 30; Mismatches 88; Indels 43; Gaps 7;

QY 24 TISSNNVIQLTKLDSNGNPVSTSVGRVLYSAPRLWFESSTVSTFTFTFQISTPYTSP 83
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 745 TESSAPVSTSTTESSAPVPTPSSSTTES-----SSAPVPTPSSSTTESSAPVPTP 797

QY 84 PGDGLAFFLPADYPIVPNSAGNI LGLFPNLNLRNSTTSKETTTIDVNAASNVAVEFD 143
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 798 SSSITTESSAVPVPT---PSSSNITSSAPSPSTPSSSTTESSSVPVPTPSSS-----TTSS 850

QY 144 TYPNDNIGDPYRKHIGDIVNSIRSKATVA-----WDNQCKTATAHSYNASKRLSVT 197
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 851 SAP-----VSSSTTESSAVPVPTPSSSNITSSAPSPSPSTTESFG 895

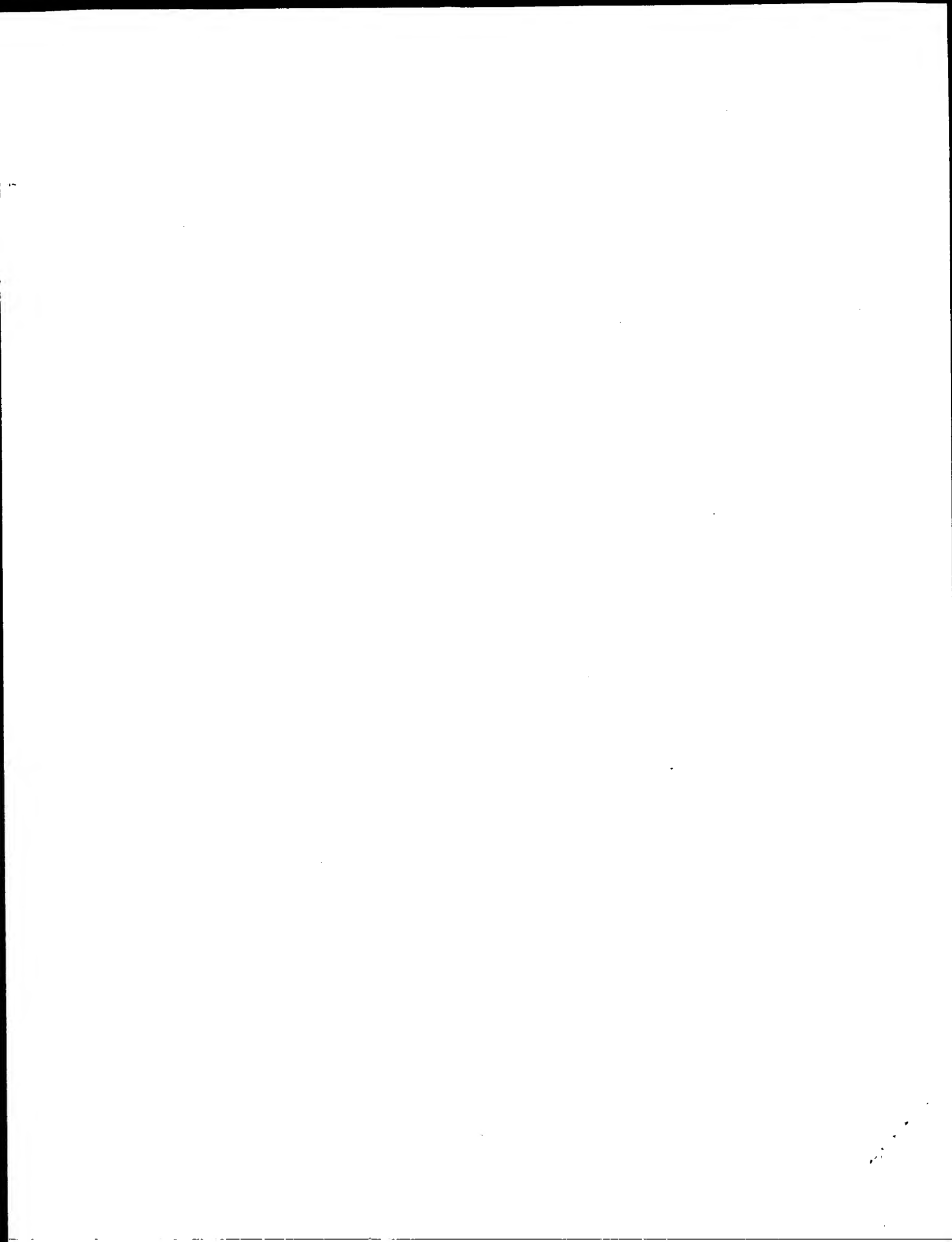
QY 198 TF-----YPGGKA-VLSLHDVELTVQ 217
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 896 TTVPSSSKYPGSQTETSVSSTTTETTIV 923

Wed Feb. 26 17:04:16 2003

us-09-476-485a-8.rapb

Page 7

Search completed: February 26, 2003, 16:52:43
Job time : 17.1798 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 15 5094 Seconds
(without alignments)
574.823 Million cell updates/sec

Title: US-09-476-485A-6
Perfect score: 1599
Sequence: 1 AOSLSFNTKFDLDQKDLIF.....LNNHKYVRCSTCMFMKKK 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCrUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	811.5	50.8	270	4	US-08-881-189B-2
2	804	50.3	286	4	US-08-881-189B-23
3	493.5	30.9	285	3	US-09-141-821-1
4	489.5	30.6	285	3	US-09-141-821-4
5	485.5	30.4	285	3	US-09-141-821-2
6	483.5	30.2	285	3	US-09-141-821-5
7	402	25.1	132	4	US-08-881-189B-12
8	383	24.0	262	2	US-08-038-761A-1
9	308.5	19.3	274	3	US-09-141-821-3
10	267.5	16.7	105	4	US-08-881-189B-13
11	221.5	13.9	632	4	US-09-228-986-77
12	107	6.7	679	4	US-08-913-942-15
13	107	6.7	679	4	US-09-268-347-26
14	101	6.3	1375	4	US-09-210-361-4
15	101	6.3	1375	4	US-09-740-274-4
16	100	6.3	852	1	US-08-471-033-36
17	100	6.3	852	2	US-08-471-044-36
18	100	6.3	852	2	US-08-463-483A-36
19	100	6.3	852	2	US-08-471-046A-36
20	100	6.3	852	2	US-08-470-566B-36
21	100	6.3	852	2	US-08-469-334-36
22	100	6.3	852	3	US-09-300-529-36
23	100	6.3	881	4	US-08-960-780-32
24	100	6.3	881	4	US-09-073-898-32
25	100	6.3	884	1	US-08-471-033-5
26	100	6.3	884	2	US-08-471-044-5
27	100	6.3	884	2	US-08-463-483A-5

28 100 6.3 884 2 US-08-471-046A-5 Sequence 5, Appli
29 100 6.3 884 2 US-08-470-566B-5 Sequence 5, Appli
30 100 6.3 884 2 US-08-469-334-5 Sequence 5, Appli
31 100 6.3 884 3 US-09-300-529-5 Sequence 5, Appli
32 100 6.3 1338 1 US-08-471-033-50 Sequence 50, Appli
33 100 6.3 1338 2 US-08-471-044-50 Sequence 50, Appli
34 100 6.3 1338 2 US-08-463-483A-50 Sequence 50, Appli
35 100 6.3 1338 2 US-08-471-046A-50 Sequence 50, Appli
36 100 6.3 1338 2 US-08-470-566B-50 Sequence 50, Appli
37 100 6.3 1338 2 US-08-469-334-50 Sequence 50, Appli
38 100 6.3 1338 3 US-09-300-529-50 Sequence 50, Appli
39 100 6.3 1346 1 US-08-471-033-23 Sequence 23, Appli
40 100 6.3 1346 2 US-08-471-044-23 Sequence 23, Appli
41 100 6.3 1346 2 US-08-463-483A-23 Sequence 23, Appli
42 100 6.3 1346 2 US-08-471-046A-23 Sequence 23, Appli
43 100 6.3 1346 2 US-08-470-566B-23 Sequence 23, Appli
44 100 6.3 1346 2 US-08-469-334-23 Sequence 23, Appli
45 100 6.3 1346 3 US-09-300-529-23 Sequence 23, Appli

ALIGNMENTS

RESULT 1
US-08-881-189B-2
; Sequence 2, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-2

Query Match 50.8%; Score 811.5; DB 4; Length 270;
Best Local Similarity 60.6%; Pred. No. 3.1e-75;
Matches 166; Conservative 39; Mismatches 62; Indels 7; Gaps 5;
QY 1 AOSLSFNTKFDLDQKDLIFQGDATSTNNVLQTKLDSGNGVSGVYLSAPFLWE 60
|||||:||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 1 AOSLSFNTKFDLPNOBDLIFQGHATSTNNVLQTKLDSAGNPVSSAGVLYSAPFLWE 60


```
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-4

Query Match          30.6%; Score 489.5; DB 3; Length 285;
Best Local Similarity 43.6%; Pred. No. 4.1e-42;
Matches 113; Conservative 37; Mismatches 86; Indels 23; Gaps 8;

QY 3 SLSFNFTKFDLDKDLIFQGDATSTN-NVLQTLKLDGCGNPVGASVGRVLSAPFHLWEN 61
   ||||| ||| :| ||||| ||| :| ||||| ||| :| ||||| ||| :| ||||| ||| :|
Db 34 SLSFSPKFPAPNQPYLIFQDALVTSTGVQLTNV-VNGVPSRKSIGRALYAAPFQIWDN 92
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 62 SMA-VSSFEETNLTIQISTPHYYAAGDGAFFLAPHDVTIIPNSWGKFLGLYSNVFRNSPT 120
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 93 TTGNVASFVTSFSFIQAPNATADGLAFLAPVDT--QPGDLGGLGIFKD----- 143
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 121 SENQSGDVNTDSRVVAVEFTFPNANIDPNYRHIGIDVNSIKSETARWONGKTATA 180
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 144 -----GSYNKSNQIVAVEEDTFNSHIFDPKGRHMGINVNSIVSKTVPWNWINGEVANV 197
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 181 RISYNSAKSKSTVTTFYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 198 FISYEASTKSLNASLVYPSLETSTFIIHAIIVDKVLPWVRFGFSATGIDTGYVQTNV 257
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 237 ISWFTSSSLK-NNEVKEPK 254
   ||||| ||| :| ||| |||
Db 258 LSWSFESNLPGGNSVASVK 276

RESULT 5
US-09-141-821-2
; Sequence 2, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Suszanna Magdalena BARDOCI
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELIS
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-2

Query Match          30.4%; Score 485.5; DB 3; Length 285;
Best Local Similarity 44.8%; Pred. No. 1e-41;
Matches 116; Conservative 39; Mismatches 81; Indels 23; Gaps 9;

QY 3 SLSFNFTKFDLDKDLIFQGDATSTN-NVLQTLKLDGCGNPVGASVGRVLSAPFHLWEN 61
   ||||| ||| :| ||||| ||| :| ||||| ||| :| ||||| ||| :| ||||| ||| :|
Db 34 SLSFSPKFKHSQDILFQSDALVTSGVLQTLTNDG--RVVDSIGRVLAAAPFQIWDN 91
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 62 SMA-VSSFEETNLTIQISTPHYYAAGDGAFFLAPHDVTIIPNSWGKFLGLYSNVFRNSPT 120
   :||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 92 TTGNVASFVTSFSFIKAPNEGKTAGLVEFLAPVGT-QPLKGGGLLGLF----- 141
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 121 SENQSGDVNTDSRVVAVEFTFPNANIDPNYRHIGIDVNSIKSETARWONGKTATA 180
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 142 -KDESY--NKSQIVAVEEDTFNVAWDPNGIHMGIDVNSIQSVRTVRWDWANGEVANV 197
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 181 RISYNSAKSKSTVTTFYPGME-VVALSHDVLHAELPEWVRVGLSASTG---EEKQNTI 236
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
```



```

RESULT 9
US-09-141-821-3
; Sequence 3, Application US/09141821
; Patent No. 6110891
; GENERAL INFORMATION:
; APPLICANT: Arpad Janos PUSZTAI
; APPLICANT: Szuzsanna Magdolna BARDOCZ
; APPLICANT: Richard Michael John PALMER
; APPLICANT: Neil William FISH
; APPLICANT: Gyorgy J. KOTELIS
; TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
; FILE REFERENCE: 48656
; CURRENT APPLICATION NUMBER: US/09/141,821
; CURRENT FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Human
US-09-141-821-3

```

Query Match	19.3%;	Score	308.5;	DB	3;	Length	274;
Best Local Similarity	34.7%;	Pred.	No. 1.5e-23;				
Matches	86;	Conservative	40;	Mismatches	97;	Indels	25;
Gaps							
QY	2	QSLSFNFRTFDLDQKDLIFOGDAT-STNNVNLQTLKDSGGNPVGASVGRVLSPASPFHLWE	60				
Db	31	EGISFNFTNTRCDQGVTLGQANIMANGILALT---NTNPTW-NGRRLYSKVPFIWD	86				
QY	61	NSMA-VSSPETNLTIOISTPHYYAAGFAFFLAPHDTVIPPNMSGKFLGLYSNVFRNSP	119				
Db	87	SATGNVASFVTSFVYKIKGGIPADGIVFFLAP-EARIPDNASGGQLGI-----	136				
QY	120	TSENQFGDGYNTDSRVAVVEFTFPNANIDPNVRHIGIDVNSIKSKETARWONGKGTAT	179				
Db	137	VNANKAY-----NPFVGVEFDIYSN-NDPKSAHIGIDASSLISLRTVYKWNKVSGLVK	189				
QY	180	ARTSYNSAKSKSVITTFYFGMEVVALSHDVLHAELPEWVRGLSA--STGEKOKNTII	237				
Db	190	VSIYDLSKLTUSVVVTHENGQISTIAQVVDLKAVLGEKVRGFTAAATTTGRIVELYDIH	249				
QY	238	SWSFTSSL	245				
Db	250	AMSFTSTL	257				

```

RESULT 10
US-08-881-189B-13
; Sequence 13, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Felt, Irving
: REGISTRATION NUMBER:
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 8
: TELEFAX: (516) 822
: INFORMATION FOR SEQ ID
: SEQUENCE CHARACTERISTICS
: LENGTH: 105 amino
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: pepti
US-08-881-189B-13

```

Query Match	16.7%;	Score 267.5;	DB 4;	Length 105;
Best Local Similarity	55.8%;	Pred. No. 5.3e-20;		
Matches 63;	Conservative 14;	Mismatches 23;	Indels 13;	Gaps 4;

QY		1	AQSLSFNFTEPDDLDKDLIFOGDNTSTNNVLQLTKLDSGGNPVGASVGRVLFSAPFHLWE	60
Db		1	: : : : : : : : : : :	
Db		1	AQSLSFSETFDPNQEDLLIFOGTAT-----KLDSAGNPVSASSAGRVLYSAPLRLEW	52
QY		61	NMSAVSSFETNLTIQTISTPHPIYYAAGDFAFFLAPHDVTIPPNSWGKFGLGYSN	113
Db		53	: : : : : : : : : : : : :	
Db		53	DSAVLTSEDP--TVITFTNTSRIADGLA-FIAPPDSVI---SYHGFGFGLGPEN	100

```

RESULT 11
US-09-228-986-77
; Sequence 77, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Niels
; TITLE OF INVENTION: Compositions Isolated
; TITLE OF INVENTION: and Their Use in the
; FILE REFERENCE: 11000/1020
; CURRENT APPLICATION NUMBER: US/09/228,986
; CURRENT FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 77
; LENGTH: 632
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-228-986-77

```

```

Query Match      13.9%; Score 221.5; DB 4; Length 632;
Best Local Similarity 27.9%; Pred. No. 5e-14;
Matches 84; Conservative 50; Mismatches 106; Indels 61; Gaps 14;

  QY      6 FNETKFDLKDGLIFQCDAT--STNNVLQITKLDGGNPGVAGSRVLFSAPFHLWNSM 63
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
  Db      40 FNGT-----NLILEANASVIGESVLSLT-----NHSHEFMLGRALYAAPQVM-KNNH 86
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

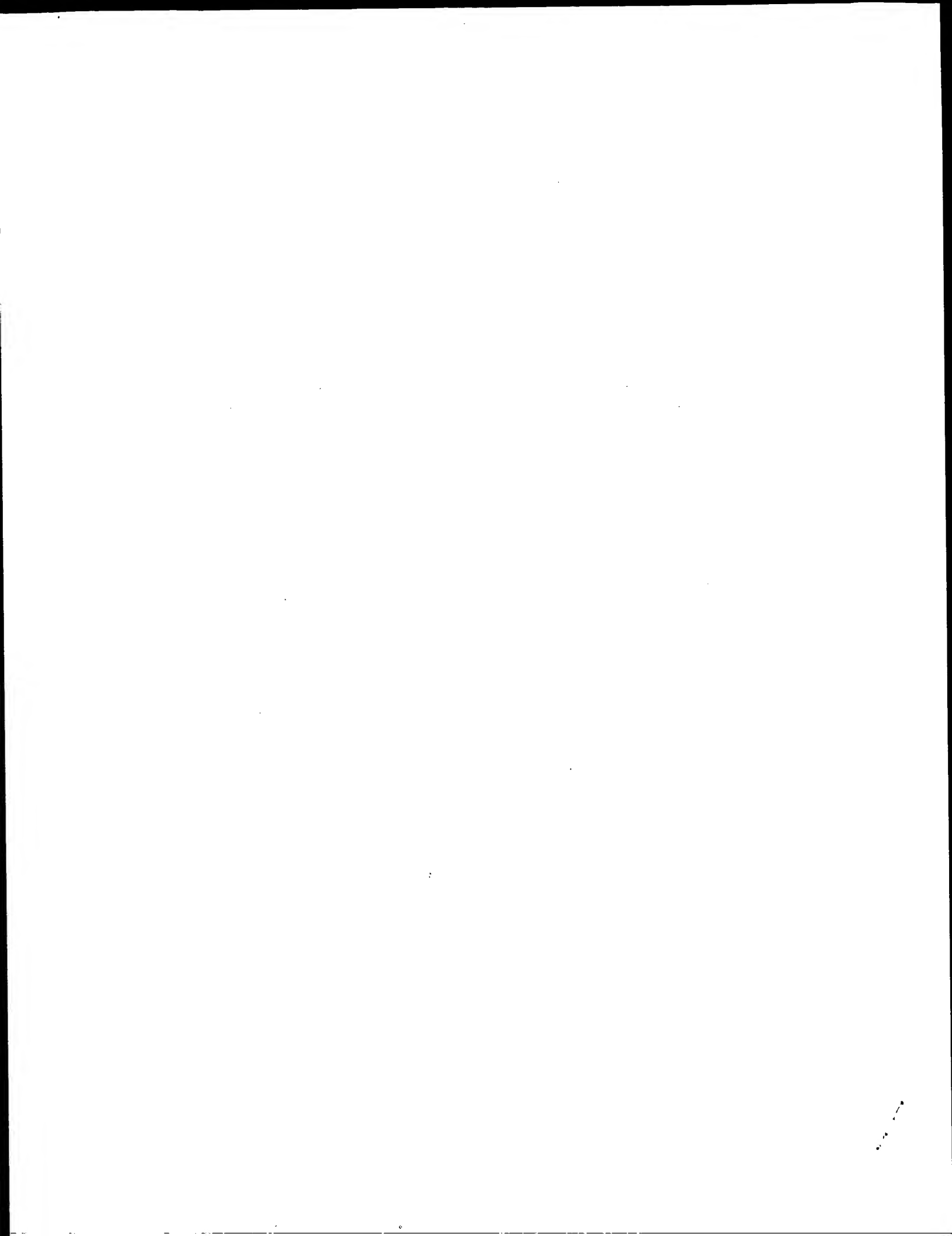
  QY      64 AVSSEFTNLTIQTSTPHYYAAGDAFAFLAPHDPTVIPPNSWGKFLGLYSNVFRNSPTSEN 123
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db      87 TVSFSFTTFVSFLVPPSPNCGGHLGALIMTPYTPSPMGAAQP-VOYLGLL-----NLTSG 139
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY      124 QSFQGDNTDSRVVAVEFDTEPPNANI-DPNYRHIGIDVNSIKS--KETARMEW----- 172
      ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  Db      140 QPY-----NHLFAVEFDTIMNVEFKDPDRNHVGVCDINSLISVQETAGY-WNGEERFEL 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY      173 --QNGTATARISYNSASKSTVTFYPGM---EWVALSHDVLDAELPEWVRVGLGSAST 227
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
  Db      193 NLRSGRNIQAWIDYDHLESSLNTVITVAGLPRQRPRLISLQIDLQNIVEKMLGVFSAAT 252
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

  QY      228 GEEKQKNTIISWFS-----SLKNEVKEPEKEDMYTANVRSYTVTINDVLSY 275
      | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
  Db      253 GLLVDEHYIWLAVSFTEETDAPPLDVCSLSSFANNYSPLSRGFTAGY----TVSVVLFV 308
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 : Search time 10.5918 Seconds
(without alignments)
1079.114 Million cell updates/sec

Title: US-09-476-485A-6

Perfect score: 1599

Sequence: 1 AOSLSFNFTKFDLDQKDLIF.....LNNHKYVRCTCMLFMKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	202.5	12.7	669	10	US-09-862-027-27
2	101	6.3	1375	10	US-09-740-274-4
3	100	6.3	881	10	US-09-850-351A-32
4	97	6.1	3092	10	US-09-967-805-7
5	96.5	6.0	784	10	US-09-801-368-172
6	92.5	5.8	1024	10	US-09-879-257A-30
7	91.5	5.7	730	10	US-09-801-368-126
8	91	5.7	1250	10	US-09-801-368-364
9	89	5.6	425	10	US-09-850-351A-27
10	88	5.5	2042	9	US-10-192-584-6
11	87.5	5.5	234	10	US-09-815-242-5763
12	87.5	5.5	268	10	US-09-815-242-12678
13	87	5.4	632	10	US-09-853-533A-8
14	86.5	5.4	2383	10	US-09-912-020-302
15	85.5	5.3	589	10	US-09-972-086-2
16	85.5	5.3	1475	10	US-09-740-274-2
17	84	5.3	1066	9	US-09-423-126-3
18	84	5.3	1066	9	US-09-280-197-5
19	83.5	5.2	1198	10	US-09-866-582-36

20	83.5	5.2	1241	12	US-10-001-215-5	Sequence 5, Appli
21	82.5	5.2	552	10	US-09-909-464A-2	Sequence 2, Appli
22	82.5	5.2	555	9	US-10-121-032-24	Sequence 24, Appli
23	82.5	5.2	2076	10	US-09-815-242-5815	Sequence 5815, Ap
24	82.5	5.2	2186	10	US-09-815-242-12913	Sequence 2, Appli
25	82.5	5.2	3014	10	US-09-737-149-25	Sequence 25, Appli
26	82.5	5.2	3034	10	US-09-737-149-30	Sequence 30, Appli
27	82.5	5.2	3034	10	US-09-737-149-30	Sequence 42, Appli
28	82	5.1	792	9	US-10-055-364-42	Sequence 4, Appli
29	81	5.1	1316	9	US-10-120-544A-4	Sequence 12610, A
30	81	5.1	5795	10	US-09-815-242-12610	Sequence 2, Appli
31	80.5	5.0	473	9	US-09-991-053-2	Sequence 36, Appli
32	80.5	5.0	589	10	US-09-866-510-12	Sequence 2, Appli
33	80.5	5.0	1089	9	US-09-955-363-36	Sequence 90, Appli
34	80.5	5.0	1089	10	US-09-769-987-2	Sequence 4, Appli
35	80.5	5.0	1089	10	US-09-919-497-90	Sequence 6, Appli
36	80.5	5.0	1089	10	US-09-866-510-2	Sequence 8, Appli
37	80.5	5.0	1089	10	US-09-866-510-4	Sequence 10, Appli
38	80.5	5.0	1089	10	US-09-866-510-6	Sequence 8, Appli
39	80.5	5.0	1089	10	US-09-866-510-8	Sequence 13446, A
40	80.5	5.0	1089	10	US-09-866-510-10	Sequence 2, Appli
41	80.5	5.0	2001	9	US-10-072-621-8	Sequence 2, Appli
42	80	5.0	378	10	US-09-815-242-13445	Sequence 2, Appli
43	79.5	5.0	518	9	US-09-976-297-2	Sequence 220, App
44	79.5	5.0	630	10	US-09-801-368-220	Sequence 2, Appli
45	79.5	5.0	716	10	US-09-853-533A-2	

ALIGNMENTS

RESULT 1

US-09-862-027-27
; Sequence 27, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862, 027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-862-027-27

Query Match	12.7%	Score 202.5;	DB 10;	Length 669;
Best Local Similarity	24.8%	Pred. No. 3.4e-11;		
Matches	80;	Conservative	63;	Mismatches 119;
				Indels 61;
				Gaps 14;
QY	6	FNFTKFDLDQKDLIFQGGDATSTNNVLQTLKLDGSGNPGVGSYGRVLFSAFPHLWNSM--	63	
DB	28	FGNGYLYDNGSIAI---TNSKGLMKLT-----NSSEFSYGHVFNYSFVR-FKNSPNG	76	
QY	64	AVSFFETNLTQISTPHPYAAGDGAFFLAPHDTVIPPNSNGKFLGLYSNFRNPTSSEN	123	
DB	77	TVSFSFTTFVFAIVSNVNALDGHGLAFVISP-TKLPLYSRSSOYLGLF-NTLNNGDPS--	132	
QY	124	QSGFDVNTDSRVVAVEFDETPNANIDP-NYRHIGIDVNSIKSKETARWEQ-----	173	
DB	133	-----NHIVAEEDTFQNEFDDMDNNHVGIDINLSLSEKASTAGYEDDDGTFTKN	183	
QY	174	---NGKTATARIYSNASKKSTVT---TFYPGMEVVALSHDVLHAEPEWVRVGLAS	226	
DB	184	IRLINQKPIQAWIEYDSSRRQLNVTIHPILPKPIPLSLTKDLSPLYFDSNMYGFTSA	243	
QY	227	TGEEKQNTIISWTF--TSSLKNNEV-----KEPKEDMTIANVVRSYTWTINDVLSYISNKM	280	

Db 244 TGRLRSSHYILGTFKNGTASNIDISRLPKLPDRSR--STSVKKILAIISLTLAILV 301
 QY 281 YDALNNHKKVRCSTCMFPMKKK 303
 Db 302 FLTI-----SYMLFLKRRK 314

RESULT 2

US-09-740-274-4
 ; Sequence 4, Application US/09740274
 ; Patent No. US20020031826A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nichols, Scott E.
 ; TITLE OF INVENTION: Glucan-containing Compositions and Paper
 ; FILE REFERENCE: 0357CRD
 ; CURRENT APPLICATION NUMBER: US/09/740,274
 ; CURRENT FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 09/210,361
 ; PRIOR FILING DATE: 1998-12-11
 ; PRIOR APPLICATION NUMBER: 09/007,999
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/478,704
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/009,620
 ; PRIOR FILING DATE: 1998-01-20
 ; PRIOR APPLICATION NUMBER: 08/485,243
 ; PRIOR FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: 09/008,172
 ; PRIOR FILING DATE: 1998-01-16
 ; PRIOR APPLICATION NUMBER: 08/482,711
 ; PRIOR FILING DATE: 1995-06-07
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1375
 ; TYPE: PRT
 ; ORGANISM: streptococcus mutans
 ; US-09-740-274-4

Query Match 6.3%; Score 101; DB 10; Length 1375;
 Best Local Similarity 23.6%; Pred. No. 0.51;
 Matches 70; Conservative 42; Mismatches 104; Indels 80; Gaps 17;
 QY 48 GRVLFSAFHLWNSMAVSSEFTNTIQISTP-----HPYYAD----GFAFFLA--- 93
 Db 385 GALLYS-----NNSKLSQANSRYRLNRTPTNQTGKDPRTADRTGGYEFFLLANDV 438
 QY 94 -PHDTVIPPN--SWGKELGLYSNVFNPSPTSENQSGFDVNTDS--RVVAVEFDTFPNANID 149
 Db 439 DNSNPVQAEQLNLWHLFLLMFGNIYANDP-----DANFDSIRVDAYD-----NVD 483
 QY 150 PYNRHIGIDVNSIKSKETARWEQNGKTATARISYNSA--SKSTVTTFYPGMEVV----- 203
 Db 484 ADLLQIAGDY----LKAAGTKHKNDKAANDHLSILEAWSYNDTPYLHDDGDNKNNMNDNR 538
 QY 204 -----ALSHDVDLHAELPEWVRVGLSASTGEEKKNTIISSEFTSSLLKNNEVKPEKED 256
 Db 539 IRLSLYSLAKPLNQRSGNPLITNSLVNRDNNATAAVPSYSFTRA--HDSVQD----- 593
 QY 257 MYIANVRSYTWIN--DVLSY-----ISNKMVDALNNHKKVRCSTCMFLF 299
 Db 594 -LIRNIIR--TEINPNVGVSYFTTEIKKAFETYNK--DLLATEKKYTHYNTALS 644

RESULT 3

US-09-850-351A-32
 ; Sequence 32, Application US/09850351A
 ; Patent No. US2002010080A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Feitelson, Jerald S.
 ; Schnepf, H. Ernest
 ; Narva, Kenneth E.
 ; Stockhoff, Brian A.

; Schmeits, James
 ; Loewer, David
 ; Dullum, Charles Joseph
 ; Muller-Cohn, Judy
 ; Stamp, Lisa
 ; Morrill, George
 ; TITLE OF INVENTION: No. US2002010080A1 Pesticidal Toxins and Nucleotide
 ; Sequences Which Encode These Toxins
 ; NUMBER OF SEQUENCES: 144
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
 ; STREET: 2421 N.W. 41st Street, Suite A-1
 ; CITY: Gainesville
 ; STATE: FL
 ; COUNTRY: US
 ; ZIP: 32606-6669
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/850,351A
 ; FILING DATE: 07-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 09/073,898
 ; FILING DATE: 06-MAY-1998
 ; APPLICATION NUMBER: US 08/960,780
 ; FILING DATE: 30-OCT-1997
 ; APPLICATION NUMBER: US 60/029,848
 ; FILING DATE: 30-OCT-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sanders, Jay M.
 ; REGISTRATION NUMBER: 39,355
 ; REFERENCE/DOCKET NUMBER: MA-708CD1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 352-375-8100
 ; TELEFAX: 352-372-5800
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 881 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; INDIVIDUAL ISOLATE: PSI77C8
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 32:
 ; US-09-850-351A-32

Query Match 6.3%; Score 100; DB 10; Length 881;
 Best Local Similarity 21.0%; Pred. No. 0.34; 133; Indels 80; Gaps 13;
 Matches 66; Conservative 36; Mismatches 133; Indels 80; Gaps 13;
 QY 9 TKFDLDOKDLIFQGDATSTNNVLQTLKLDGNNP-----VGASGRVLFSAF- 55
 Db 157 TKFNIDSK-----TFKELKLFKIDSONQPOVOQDELNPFNKKESQEFLLAKPS 206
 QY 56 -FHLWNSMAVSSEFTNTIQTISPHYPYAADGFAF---FLAPHDTVIPPNWSGKFLGLY 111
 Db 207 KINLFOTKMKREIDEDTDTGDSIP-DLWEENGYTIONRIAVKWDSDLSASKYTKFV--- 262
 QY 112 SNVFNPSPTSENQSGFDVNTDSRVVAVEFD-----TFPNANI-----DPN 151
 Db 263 -----SNPLESHVTGDPYTDYKAAARDLDSNAKETNPVLAAPFVSNVSMKVLSNP 316
 QY 152 YRHIGIDVNSIKSKETARWEQNGKTATARISYNSASKSTVTTFYPGMEVVVALSHDVDL 211
 Db 317 ENL-----SNSVESHSSTNWSYNTNEGASVEAGIGPKGISFGVSVNYQHSSETVA----- 365
 QY 212 HAELEPWVRVGLSASTGEEKKNTIISSEFTSSLLKNNEVKPEKEDMYTANVRSYTWIND 271

```

366  Db  -----QEW-----GTSGTNSQFN TASGLYLNANVRNNVTGA--IYDYKPTTSEVLNND 414
272  QY  VLSYISNKMYD-ALN 285
      ::  |
      ::  |
415  Db  TIATITAKSNSTALN 429

```

RESULT 4
US-09-801-368-172
; Sequence 172, Application US/09801368
; Patent No. US20020128250A1

```

; WORKING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Peptide sequence
US-09-967-805-7

```

Query Match 6.0%; Score 96.5; DB 10; Length 784;
Best Local Similarity 21.3%; Pred. No. 0.63;
Matches 63; Conservative 35; Mismatches 107; Indels 91; Gaps

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

```

366  Db  -----QEW-----GTSGTNSQFN TASGLYLNANVRNNVTGA--IYDYKPTTSEVLNND 414
272  QY  VLSYISNKMYD-ALN 285
      ::  |
      ::  |
415  Db  TIATITAKSNSTALN 429

```

RESULT 4
US-09-801-368-172
; Sequence 172, Application US/09801368
; Patent No. US20020128250A1

```

; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary

```

; APPLICANT: No. US20020128250Alman, Thea
 ; APPLICANT: Royer, John
 ; APPLICANT: Salama, Sofie

TITLE OF INVENTION: NO. US2002010080A1el Pesticidal Toxins and Nucleotide Sequences Which Encode These Toxins

NUMBER OF SEQUENCES: 144

CORRESPONDENCE ADDRESS:

ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: US

ZIP: 32606-6669

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/850,351A

FILING DATE: 07-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/073,898

FILING DATE: 06-May-1998

APPLICATION NUMBER: US 08/960,780

FILING DATE: 30-OCT-1997

APPLICATION NUMBER: US 60/029,848

FILING DATE: 30-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Sanders, Jay M.

REGISTRATION NUMBER: 39,355

REFERENCE/DOCKET NUMBER: MA-708CD1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 352-375-8100

TELEFAX: 352-372-5800

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 425 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: 68F

SEQUENCE DESCRIPTION: SEQ ID NO: 27:

US-09-850-351A-27

Query Match 5.6% Score 89; DB 10; Length 425;

Best Local Similarity 20.68; Pred. No. 1.4; Indels 90; Gaps 12;

Matches 59; Conservative 31; Mismatches 106; Indels 90; Gaps 12;

QY 20 FQGDATSTNNVQLTK-----LDSGGNPVGASVGRVLFSAFPHLWENSMAYSSPETNLT 73

Db 152 FLTKATKTNLTQVKSTRDEDTDGD-----SIP-DIWE-----ENGYT 191

QY 74 IQISTPHY---YAADGFAPFLA-PHDTVIPPNMGKFLGLYSNVFRNSPTSENSGFGDV 129

Db 192 IQNKIAVKWDSLASKGYTKVSNPLDT-----HTVGGP 225

QY 130 NTDSRVVAFED-----TFPNANIDPNYRHIGID---VNSIKKETARWEQ 173

Db 226 YTDEKARDLDLSNAKETFNPLVAAPFSVNVSMKEVILSPDENLSNIESHSSTNWST 285

QY 174 NGKTATARISYNSASKSTVTTFYPGMEVVALSHDVLHAEPEWVRVGLSASTGEKQK 233

Db 286 NTEGASIEAGGALGFGVSANTQHSETVGY-----EW-----GTSTGNTSGF 329

QY 234 NTIISWFTSLKKNVEKPEKEDMYIANVRSYTWINDVLSYISNK 279

Db 330 NTASAGYLNANVRYNVGTGA--IYDVKPTTSFVLNKKDTIATITAK 373

RESULT 10

US-10-192-584-6

; Sequence 6, Application US/10192584

; Publication No. US20030027987A1

; GENERAL INFORMATION:

; APPLICANT: TOKUNAGA, Eiji

; SAKAGUCHI, Masashi

; MATSUO, Kazuo

; HAMADA, Fukuaburo

; TOKIYOSHI, Sachio

; TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 624 Ninth Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/192,584

; FILING DATE: 11-Jul-2002

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/077,098

; FILING DATE: 19-May-1998

; APPLICATION NUMBER: PCT/JP97/03222

; FILING DATE: 12-SEP-1997

; APPLICATION NUMBER: JP 27,148/1996

; FILING DATE: 19-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: KORNBAU, Anne M.

; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: TOKUNAGA-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-628-5197

; TELEFAX: 202-737-3528

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2042 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-192-584-6

Query Match 5.5% Score 88; DB 9; Length 2042;

Best Local Similarity 22.4%; Pred. No. 16;

Matches 62; Conservative 42; Mismatches 107; Indels 66; Gaps 13;

QY 22 GDATS---TNVLQKLKDSGGNPVGASVGRVLFSAFPHLWENSMAYSSFETNLTQIST 78

Db 1209 GDANNKQLNNTLQQTLEATG--ITSSVGSNTY-AGFSLGADSVTFSGGAG-TVKLSG 1264

QY 79 PHPYAAADGFA-----FFLAPHDTVIPPNWSGKFLGLYSNVFRNSPTSENQSGFD 128

Db 1265 VSDATADTAATLKQVKRYRTTLVGDNDITAADRSGGTSNGITYNLSLNGKTVS----- 1318

QY 129 VNTDSRVVA--VEFDTFPNANIDPNYRHIGID---VNSIKKETARWQNGKTATARIS 183

Db 1319 -ATEEKVVSQKTVYEAIRNA-ITGNIFTGLDDTTLNKNIN-----PADQDLS 1364

QY 184 YNSASKKSTVTTFYPGMEVVALSHDVLHAEPEWVRVGLSASTGEKQKQKTIISWSETS 243

Db 1365 NLSESGKNAITGL--VDVVVKINS-----PITVEPSTDSNKKKTFVGVDFD 1410

QY 244 SLKNNVEKPEKEDMYIANVRSYTWINDVLSYISNM 280

Db 1411 TITEGDATDDK-----LTTSKSVESYVTNKL 1437

RESULT 11

US-09-815-242-5763
; Sequence 5763, Application US/09815242
; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 5763

; LENGTH: 234

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-5763

Query Match

Best Local Similarity 5.5%; Score 87.5; DB 10; Length 234;

Mismatches 51; Conservative 26; Mismatches 79; Indels 59; Gaps 11;

QY 9 TKFDLDQKDLIFQGDATSTNNVLQTLKDSGPNVGASVGRVLFSAFFHLWENSMVSSSF 68

Db 4 TNLNVRDLDM--TTFYKNILGLSVKSSDDNTTVLSVGTGGHTLTLLHLEDGRTSPR 61

QY 69 ETNL-TIQISTPHPYAAGDPAFFLAPHDTVIPPNWSGKFLGLYSNFRNSPTSENQSGF 127

Db 62 EAGLFHIAFLPTTDLAN-FLYFVAQ-----KNMGIGAG---DHLVSEALYFN 106

QY 128 DVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSETARWONGKTATARISYNSA 187

Db 107 D-----PEGN-----GIEV---YDRPSSSWWONGKVKMDTLEVDS- 140

QY 188 SKKSTVTTF-----YPGMEVVAL-----SHDVD 210

Db 141 ---QTLTHRTDEGWQGMKPAKGIMHGLHKLKTHDLD 172

RESULT 12

US-09-815-242-12678

; Sequence 12678, Application US/09815242

; Patent No. US20020061569A1

; GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.

; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO: 12678

; LENGTH: 268

; TYPE: PRT

; ORGANISM: Staphylococcus aureus

US-09-815-242-12678

Query Match 5.5%; Score 87.5; DB 10; Length 268;

Best Local Similarity 23.7%; Pred. No. 1;

Mismatches 51; Conservative 26; Mismatches 79; Indels 59; Gaps 11;

QY 9 TKFDLDQKDLIFQGDATSTNNVLQTLKDSGPNVGASVGRVLFSAFFHLWENSMVSSSF 68

Db 12 TNLNVRDLDM--TTFYKNILGLSVKSSDDNTTVLSVGTGGHTLTLLHLEDGRTSPR 69

QY 69 ETNL-TIQISTPHPYAAGDPAFFLAPHDTVIPPNWSGKFLGLYSNFRNSPTSENQSGF 127

Db 70 EAGLFHIAFLPTTDLAN-FLYFVAQ-----KNMGIGAG---DHLVSEALYFN 114

QY 128 DVNTDSRVVAVEFDTPFNANIDPNYRHIGIDVNSIKSETARWONGKTATARISYNSA 187

Db 115 D-----PEGN-----GIEV---YDRPSSSWWONGKVKMDTLEVDS- 148

QY 188 SKKSTVTTF-----YPGMEVVAL-----SHDVD 210

Db 149 ---QTLTHRTDEGWQGMKPAKGIMHGLHKLKTHDLD 180

RESULT 13

US-09-853-533A-8

; Sequence 8, Application US/09853533A

; Patent No. US20020103362A1

; GENERAL INFORMATION:

; APPLICANT: Monsanto Company

; APPLICANT: Isaac, Barbara

; APPLICANT: Krieger, Elysia

; APPLICANT: Mettius, Anne-Marie

; APPLICANT: Moshiri, Farhad

; APPLICANT: Sivasubramanian, Sakuntala

; TITLE OF INVENTION: POLYPEPTIDE COMPOSITIONS TOXIC TO ANTHONOMUS INSECTS, AND METH

; FILE REFERENCE: 38-21(51932)B

; CURRENT APPLICATION NUMBER: US/09/853,533A

; CURRENT FILING DATE: 2001-05-11

; PRIOR APPLICATION NUMBER: 60/204,367

; PRIOR FILING DATE: 2000-05-15

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 8

; LENGTH: 632

; TYPE: PRT

; ORGANISM: Bacillus thuringiensis

```

;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28) (30)
; OTHER INFORMATION: alternative methionine initiation codon sequence
US-09-853-533A-8

Query Match      5.4%; Score 87; DB 10; Length 632;
Best Local Similarity 20.6%; Pred. No. 3.8;
Matches 64; Conservative 38; Mismatches 103; Indels 106; Gaps 17;

QY 11 FDLQKDLIFOGDAT--STNNVLQL-----TKLDGGNPNVGSVGRV---LFSAPPH 57
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 210 FDLNOKYTFIIGSATGANNKHQIGVTLFEAYFTKPTTEANPDVIELGTAFLPLNHEPIC 269
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 LWNEMASVSPETNLTIQ---ISTPHPYAAGGFAFFLAPHDTVPPNSWGKFLGLYSNV 114
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 270 LKATDEVGDITKQITVEFNIDITSKP-----GAYRTYKVV 306
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 115 FRNSPTSENOQSGDVNTDSRVVAVEEDTFPNAN-----IDP-NYRHIG----- 156
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 307 -----NSYGE--SDEKTEIVVVYTKPTITAHDIITIKKDLATFDPLNYEPIGLKATDP 355
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 157 -----IDVNSIKSKETARWQNGKTATARI--SYNSASKSKSVTTPYPGMEVVA 204
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 IDGITDKIAVKFNNDVTSKPKYH-----VTYKVINSYKIDDKTEIVTYTKPSIVA 409
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 205 LSHDVLHAEPL-----EMVRVGLSAS---TGEKOKNTIISWSFTSSLKNNVEKPKEDM 257
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 410 --HDVETKKTAFDPLNYEPIGLKATDPIDGIDITDKITV-----ESNDVDTSKPGA 458
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 258 YIA--NVRSY 266
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 459 YSVYKVVNNY 469
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
US-09-912-020-302
; Sequence 302, Application US/09912020
; Patent No. US20020045592A1
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlson, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; FILE REFERENCE: ELITRA.001DV1
; CURRENT APPLICATION NUMBER: US/09/912,020
; PRIOR FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/492,709
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 60/117,405
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 302
; LENGTH: 2383
; TYPE: PRT
; ORGANISM: E. Coli
US-09-912-020-302

Query Match      5.4%; Score 86.5; DB 10; Length 2383;
Best Local Similarity 19.3%; Pred. No. 28;
Matches 65; Conservative 50; Mismatches 125; Indels 97; Gaps 12;

QY 14 DQDKLIFOGDATSNVYLQTLKLDGGNPNVGCASVGRVLFSAFPHLWNSMAV--SSFETN 71
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1158 DSQPVTFVADKASQVVLQISKDEITGNGVDSATLTATVKDQDFONEVNNLPVTFSSSG 1217
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

QY 72 LTIQ--ISTPHPYAAD---GFAF-----FLA 93
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1218 LTLTPGVSNTNESAIAQATLAGVAFGEKTVTASLANNGASDNKTVHFTGDTAAAKIIELA 1277
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 94 P-HDTVI---PPNSWGKFL-----GLYSNVFRNSPTSENOQSGDVNTDSRVV 136
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1278 PVPDSIIAGTPQNSGSGVITATVVDNNGFPVKGVTVVFTSNAATAEWNGCAVTNEOGK 1337
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 137 AVFEDTFPNANIDPNYRHIGIDVNSIKSKETARWQNG--KTATARISYNSASKSKSVTT 195
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1338 ATVTYNTNRSIESGAR-----PDTVEASLENGSSTLSTSIINVNADASTAHLTL 1386
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 196 FYPGMEVVALSHDVLHAEPL-----PEWVRVGLSASGEEKQKNTIIS----- 238
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1387 LQALFDTVSAGETSLYIEVKDNGVGPQEQEVTLSPSEGVTPSNNAIYTTNHDGNFY 1446
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 239 -----WSTSSLKNNVEKPKEDMYIANVRS 265
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1447 ASPTATKAGVYQLTATLENGDSMQ-QVTVTVPNVANA 1482
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-09-972-086-2
; Sequence 2, Application US/09972086
; Patent No. US20020090710A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Curtis, Rory
; TITLE OF INVENTION: 57800, A NOVEL HUMAN CADHERIN AND USES
; FILE REFERENCE: 38155-20038.00
; CURRENT APPLICATION NUMBER: US/09/972,086
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/237,698
; PRIOR FILING DATE: 2000-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-972-086-2

Query Match      5.3%; Score 85.5; DB 10; Length 589;
Best Local Similarity 24.4%; Pred. No. 4.8;
Matches 63; Conservative 39; Mismatches 87; Indels 69; Gaps 16;

QY 60 ENSMAYSSFETN-----LTIQISTPHPYA-ADGFAFFLAPHDTVPPNSWGKFLGLY 111
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 154 ENSAAI-DFEKSQVLTFLKLAVERNTEPEKFSSTADVVLIQLDITNDV-----PKFDSLY 206
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 112 --SNVFRNSPTSENOQSGDVNTDSRVVAVEEDTFPNANIDPNYRHIGIDVNSIKSKE--- 166
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 207 YVARIPENAPGG-----SSVAVTAVDPDTGPMGEVKYSTYCTGADFLIHPSTGLI 258
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 167 -TARWQNGKTATARISYNSASKSKSVTTFYPCGMEVVALSHDVLHAEPLPEWVR----- 220
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 YTPQWASLDAE-ATAR--YNFYKAEDECKYSVAEFTILLVDNDHP--PQGRSVOKK 313
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 221 -----VGLSASTGEEKQKNTIISWSFTSSLKNNVEKPKEDMYIANV--VRSYT---W 268
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 314 TWVLGTPVKIEADEDAEENNLVDYSITHA-----EP-----ANVFDSHTGEIW 360
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 269 INDVLSYISNKMYDALNN 286
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 LKN-----SIRSLDALHN 373
   |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: February 26, 2003, 16:52:34
Job time : 16.5918 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 : Search time 37.0712 Seconds
(without alignments)
1089.120 Million cell updates/sec

Title: US-09-476-485A-6

Perfect score: 1599

Sequence: 1 AQLSLFNFTRFDLQDKLIF.....LNNHXYVRCSTCMFMKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1599	100.0	303	22	AAG62898
2	804	50.3	264	20	AAW87973
3	804	50.3	264	22	AAG62890
4	804	50.3	286	22	AAG62894
5	787.5	49.2	234	22	AAG62901
6	658	41.2	290	21	AAV58736
7	653	40.8	290	21	AAV58737
8	564	35.3	237	16	AAW74765
9	493.5	30.9	285	20	AAV06811
10	489.5	30.6	285	20	AAV06814

11	485.5	30.4	285	20	AAV06812	R. pseudoacacia le
12	483.5	30.2	285	20	AAV06815	Amino acid sequenc
13	471	29.5	275	21	AAV58738	Pea lectin. 2ltn.
14	416	26.0	228	16	AAW74766	Pea lectin. 2ltn.
15	383	24.0	262	20	AAV27265	B. purpurea native
16	378	23.6	242	15	AAW45911	Bauhinia purpurea l
17	373	23.3	242	15	AAW45912	Bauhinia purpurea l
18	350	21.9	236	16	AAW64807	PNA lectin subunit
19	345.5	21.6	239	16	AAW74767	Pea lectin, ille.
20	308.5	19.3	274	20	AAV06813	R. pseudoacacia le
21	287	17.9	265	10	AAV91967	Sequence of arceli
22	250.5	15.7	282	21	AAW25452	Pinus radiata cell
23	245	15.3	652	23	ABW33374	Herbicidally activ
24	243	15.2	681	23	ABW33374	Herbicidally activ
25	237.5	14.9	244	10	AAV93640	Sequence of a Phas
26	221.5	13.9	632	21	ABW25109	Pinus radiata cell
27	209.5	13.1	715	23	ABW92649	Herbicidally activ
28	203.5	12.7	283	23	ABW91344	Herbicidally activ
29	203	12.7	711	23	ABW93337	Herbicidally activ
30	202.5	12.7	669	23	ABW93342	Herbicidally activ
31	198	12.4	649	23	ABW93342	Herbicidally activ
32	197.5	12.4	675	23	ABW92029	Herbicidally activ
33	197.5	12.4	677	23	ABW92654	Herbicidally activ
34	195.5	12.2	674	23	ABW93878	Herbicidally activ
35	193.5	12.1	691	23	ABW93681	Herbicidally activ
36	188.5	11.8	272	23	ABW91343	Herbicidally activ
37	186	11.6	242	23	ABW91342	Herbicidally activ
38	185	11.6	604	23	ABW92522	Herbicidally activ
39	179.5	11.2	566	23	ABW91530	Herbicidally activ
40	177	11.1	623	23	ABW91919	Herbicidally activ
41	177	11.1	682	23	ABW92521	Herbicidally activ
42	176.5	11.0	693	23	ABW92247	Herbicidally activ
43	169.5	10.6	685	23	ABW93319	Herbicidally activ
44	167	10.4	657	23	ABW93898	Herbicidally activ
45	162.5	10.2	627	23	ABW91918	Herbicidally activ

ALIGNMENTS

RESULT 1

AAG62898
ID AAG62898 standard; Protein: 303 AA.

AC AAG62898;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a french bean FRIL polypeptide.

KW FRIL; FLK2/FLK3 Tyrosine kinase receptor-interacting lectin;
KW progenitor cell preservation factor; radiotherapy; chemotherapy;
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
KW severe combined immunodeficiency; aplastic anemia; tissue repair.

OS Phaseolus vulgaris.

PN WO200149851-A1.

PD 12-JUL-2001.

PF 30-DEC-1999; 99WO-US31307.

PR 30-DEC-1999; 99WO-US31307.

PA (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

DR WPI; 2001-441882/47.

DR N-PSDB; AAH42306.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo

PT preservation of hematopoietic progenitor cells and as therapeutics for
PT alleviating/reducing progenitor cell-depleting activity of cancer
XX therapeutics -
XX
XX Example 5; Page 81; 173pp; English.
XX
CC The present sequence represents a FRIL (FLK2/FLT3 tyrosine kinase
CC receptor-interacting lectin) polypeptide. The specification describes a
CC composition of one or more members of FRIL family of progenitor cell
CC preservation factors. The composition is useful for alleviating or
CC reducing the hematopoietic progenitor cell-depleting activity of
CC a therapeutic treatment, including radiotherapeutic and/or
CC chemotherapeutic treatments. Administration of FRIL compositions to a
CC patient prior to treatment of the patient with a therapeutic treatment
CC having a hematopoietic progenitor cell-depleting activity of the
CC reduces the hematopoietic progenitor cell-depleting activity of the
CC therapeutic treatment in the patient. FRIL family members are useful for
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
CC stem cells. The composition is administered to reduce progenitor cell
CC depleting effects of chemotherapeutics, so that the patient can receive
CC a higher dose of the chemotherapeutic and preferably recover from cancer.
CC It is also administered to patients having, or predisposed to developing
CC a condition where the patients hematopoietic progenitor cells are
CC depleted, such as severe combined immunodeficiency or aplastic anemia.
CC The isolated mesenchymal cells are useful for tissue repair.
XX
XX Sequence 303 AA;
XX
XX Query Match 100.0%; Score 1599; DB 22; Length 303;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-141; Indels 0; Gaps 0;
XX Matches 303; Conservative 0; Mismatches 0;
XX
XX 1 AQSLSFNTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60
XX Db 1 AQSLSFNTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60
XX
XX 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLYSNVRNSPT 120
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 120
XX Db 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLYSNVRNSPT 120
XX
XX 121 SENQSGDVNTDSRVVAVVEFDFFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 180
XX Db 121 SENQSGDVNTDSRVVAVVEFDFFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180
XX
XX 181 RISYNSASKKSTVTTFYPGMEVVALSHVDLHAELEPEWVRVGLSASTGEEKQKNTIISWS 240
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 240
XX Db 181 RISYNSASKKSTVTTFYPGMEVVALSHVDLHAELEPEWVRVGLSASTGEEKQKNTIISWS 240
XX
XX 241 FTSSLKNNVEKPEKEDMYIANVVRSTWINDVLSYISNKMWDALNNHKKYVRCSTCMLFM 300
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 300
XX Db 241 FTSSLKNNVEKPEKEDMYIANVVRSTWINDVLSYISNKMWDALNNHKKYVRCSTCMLFM 300
XX
XX 301 KKK 303
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 303
XX Db 301 KKK 303
XX
XX RESULT 2
XX ID AAW87973
XX AC AAW87973;
XX XX AAW87973;
XX DT 13-APR-1999 (first entry)
XX DE A lectin derived progenitor cell preservation factor.
XX KW Lectin derived progenitor cell preservation factor; progenitor cell;
XX KW haematopoietic cell; cultured cell preservation; anticancer therapy;
XX KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;
XX KW FLK2/FLT3 receptor.
XX XX Dolichos lab lab.
XX OS

XX WO9859038-A1.
XX PN 30-DEC-1998.
XX PD
XX PF 23-JUN-1998; 98WO-US13046.
XX PR 24-JUN-1997; 97US-0881189.
XX XX (IMCL-) IMCLONE SYSTEMS INC.
XX PA (REGC) UNIV CALIFORNIA.
XX XX Chrispeels MJ, Colucci MG, Moore JG;
XX PI WPI; 1999-081274/07.
XX XX N-PSDB; AAX03593.
XX DR
XX XX New nucleic acid encoding plant lectin that preserves progenitor
XX cells - particularly haematopoietic progenitors, useful for bone
XX marrow reconstitution after ablative therapy, and to increase DNA
XX transfer in gene therapy
XX PT
XX PS Claim 1; Page 30-31; 72pp; English.
XX XX
XX CC The present sequence represents a lectin derived progenitor cell
XX preservation factor. The protein is used to preserve unipotent,
XX pluripotent or totipotent progenitor cells, especially haematopoietic
XX cells, and also progenitors from nerve, muscle, skin, gut, bone,
XX kidney, liver, pancreas or thymus. Specific applications are
XX preservation of cultured cells intended for administration after
XX (anticancer) myeloablative therapy (bone marrow or whole-body irradiation
XX or chemotherapy) to reconstitute the haematopoietic system; enrichment
XX of progenitor cells (e.g. during ex vivo purging of malignant cells);
XX treatment of tissues containing haematopoietic progenitors for subsequent
XX transplant to improve haematopoietic competence; improving transfer of
XX exogenous DNA to progenitor cells (in gene therapy of various
XX haematological disorders, e.g. sickle-cell anaemia); and protection
XX against ablative therapy (to eliminate proliferating cells specifically),
XX followed by re-establishment of differentiation and proliferation of
XX preserved progenitors. The protein, when linked to magnetic beads, may
XX also be used to isolate cells that express the FLK2/FLT3 receptor.
XX
XX Sequence 264 AA;
XX
XX Query Match 50.3%; Score 804; DB 20; Length 264;
XX Best Local Similarity 60.9%; Pred. No. 6.3e-67;
XX Matches 162; Conservative 39; Mismatches 59; Indels 6; Gaps 4;
XX
XX 1 AQSLSFNTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 60
XX Db 1 AQSLSFNTKFDLDKDLIFOGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60
XX
XX 61 NSMAVSFETNLTIQISTPHYPYAADGFAFFLPHDVTIPPNSWKGFLGLY--SNVFRNS 118
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 118
XX Db 61 DSAVLTSFDTIINFESTPTYSRIADGLAFTIAPDPSVI--SYHGFGFLGFPNATLNS 118
XX
XX 119 PTSENOSEGDVNTDSRVVAVVEFDFFPNANI-DPNYRHIGIDVNSIKSKETARWQNGKT 177
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 177
XX Db 119 STSENQFTTKA-ASSNVVAVVEFDIYNPDYGDNPVNIHIGIDVNSIRSKVTAKDWQNGKI 177
XX
XX 178 ATARISYNSASKKSTVTTFYPGMEVVALSHVDLHAELEPEWVRVGLSASTGEEKQKNTII 237
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 237
XX Db 178 ATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTLVLPEWVRVGLSASTGQDKERTVH 237
XX
XX 238 SWSFTSSLKNNVEKPEKEDMYIANVV 263
XX Qy |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 263
XX Db 238 SWSFTSSLTWNVAKKENENKVIITRGV 263
XX
XX RESULT 3
XX ID AAG62890
XX ID AAG62890 standard; Protein; 264 AA.
XX XX

AC AAG62890;
 XX 17-SEP-2001 (first entry)
 XX Amino acid sequence of a hyacinth bean FRIL polypeptide.
 DE
 XX FRIL; FIK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.
 XX
 OS Dolichos lab lab.
 XX
 XX WO200149851-A1.
 XX
 XX 12-JUL-2001.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 XX (PHYL-) PHYLOGIX LLC.
 XX
 XX Colucci MG, Chrispeels MJ, Moore JG;
 XX WPI; 2001-441882/47.
 DR N-PSDB; AAH42287.
 XX
 XX Legume progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics
 XX
 PS Example 1; Page 54-55; 173pp; English.
 XX
 CC The present sequence represents a FRIL (FIK2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preserving factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of
 CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 264 AA;

Query Match 50.3%; Score 804; DB 22; Length 264;
 Best Local Similarity 60.9%; Pred. No. 6.3e-67;
 Matches 162; Conservative 39; Mismatches 59; Indels 6; Gaps 4;

QY 1 AQSLSFNFTKDLDDKDLIFQGDATSTNNVLIQTLKDSGGNPGVSGVRVLSAPFHLWE 60
 DB 1 AQSLSFSFTKFDPNQBDLIFQGHATSTNNVLIQTLKDSAGNPNVSSSAGRVLSAPLRWE 60
 QY 61 NSMAVSSFTNLTIQISTPHPYAAGDAFFAPHDPIPPNSGKFLGLY--SNVFN 118
 DB 61 DSAVLTSFTDIINFESTYTSRIAGLAFIAPPDSVI--SYHGFLGLFPNPNLNN 118
 QY 119 PRSENGSGDVNTDSRVVAFEFTFPNANI-DPNYRHIGIDVNSIKSKETARWEQNGKT 177
 DB 119 STSENQTTKA-ASSNVVAVEFTYLPDYPDNYIHIGIDVNSIRSKVTAKWDQNGKI 177

QY 178 ATARISVNSASKKSTVTTFYPCMEVVALSHDVLHAELPEWVRVGLSASTGEQKNTII 237
 DB 178 ATAHISVNSVSKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERNIVH 237
 QY 238 SWSFTSSLKKNVEKPEKEDMYIANVV 263
 DB 238 SWSFTSSLTNTVAKKENENKYYIRGV 263
 RESULT 4
 AAG62894
 ID AAG62894 standard; Protein; 286 AA.
 AC AAG62894;
 XX
 XX 17-SEP-2001 (first entry)
 XX
 XX Alpha-amylase inhibitor signal peptide and FRIL fusion.
 XX
 KW FRIL; FIK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;
 KW alpha-amylase inhibitor gene.
 XX
 XX Synthetic.
 OS Unidentified.
 OS Dolichos lab lab.
 XX
 PN WO200149851-A1.
 XX
 PD 12-JUL-2001.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 XX (PHYL-) PHYLOGIX LLC.
 XX
 XX Colucci MG, Chrispeels MJ, Moore JG;
 XX WPI; 2001-441882/47.
 DR N-PSDB; AAH42295.
 XX
 XX Legume progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics
 XX
 XX Example 1; Page 59; 173pp; English.
 XX
 CC The present sequence represents fusion protein of alpha-amylase inhibitor
 CC signal peptide and FRIL (FIK2/Flt3 tyrosine kinase receptor-interacting
 CC lectin). The specification describes a composition of one or more members
 CC of FRIL family of progenitor cell preservation factors. The composition
 CC is useful for alleviating or reducing the hematopoietic progenitor
 CC cell-depleting activity of a therapeutic treatment, including
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of
 CC FRIL compositions to a patient prior to treatment of the patient with
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting
 CC activity alleviates or reduces the hematopoietic progenitor
 CC cell-depleting activity of the therapeutic treatment in the patient.
 CC FRIL family members are useful for isolating population of progenitor
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is
 CC administered to reduce progenitor cell depleting effects of
 CC chemotherapeutics, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 286 AA;

CC The present sequence is that of the concanavalin A (Con A) lectin
 CC of jack bean. The present invention relates to: a transformed
 CC Brassica plant that is resistant to certain insect pests; an
 CC expression cassette containing DNA that codes for at least 1
 CC lectin selected from Con A, modified Con A (see AAY58737) and pea
 CC lectin (see AAY58738); transgenic plant cells containing at least 1
 CC copy of the DNA; a new lectin derived from the jack bean Con A
 CC gene; a method of imparting resistance to insects selected from
 CC blossom beetles (pollen beetles) of the genus *Meligethes*, flea
 CC beetles of the genus *Phyllotreta*, and root flies of the genus
 CC *Delia*; and a method for protecting a plant against infestation by
 CC insects of these genera.
 XX
 SQ Sequence 290 AA;

Query Match 41.28; Score 658; DB 21; Length 290;
 Best Local Similarity 52.28; Pred. No. 3.5e-53;
 Matches 140; Conservative 42; Mismatches 66; Indels 20; Gaps 9;
 QY 3 SLSEFNTKFDLQKDLIFQGDATS--TNNVLQTLKLDGNGPVGASVGRVLSAPFHLWEN 61
 DB 36 ALHFMFNQFSKQDKLILQGDATTGTGDNLELTRVSSNGSPQGSVGRALFYAPVHIWES 95
 QY 62 SMVSSFTETNLTIQISTP--HPYYAAGGFAFFLPHDTPVPNSWGKFLGLY--SNVFRN 117
 DB 96 SAVVASFEATFTFLIKSPDSHP---ADGIAPFISNIDSSIPSGSTGRLLGLFPDANVIRN 152
 QY 118 SPTSE--NQSFGDVNTDSRVVAVEFDTPPNANI--DPNVRHIGIDVNSIKSKETARWQNG 175
 DB 153 STTIDFNAAY---NADT-IVAVELDTVPNTDIGDPSPYPHIGIDIKSVRSKKTAKNNWQNG 208
 QY 176 KTATARISYNSASKKSTVTTFYFGMEVVALSHVDLHAELPEWVRVGLSASTGEEKQKNT 235
 DB 209 KVGTAHIYNSVDKRLSAVSYPNADSATVSYDVLNDVLPFWVRVGLSASTGLYKETNT 268
 QY 236 ILSWSTSSLLKNEVKEPKEDMYIANVV 263
 DB 269 ILSWSTSKLSNEIPD-----IATVV 290

RESULT 7
 AAY58737

ID AAY58737 standard; Protein: 290 AA.

XX AAY58737;

XX 25-APR-2000 (first entry)

XX Jack bean concanavalin A N152S mutant.

XX Concanavalin A; Con A; lectin; jack bean; insecticide;
 XX transgenic plant; Brassica; insect resistance; mutant; mutein.

XX Canavalia ensiformis.

XX Key Location/Qualifiers
 XX misc_difference 152

FT /note= "replaces wild-type Asn"

XX WO200001223-A1.

XX 13-JAN-2000.

XX 02-JUL-1999; 99WO-SE01209.

XX 03-JUL-1998; 98SE-0002425.

XX (PLAN-) PLANT SCI SVERIGE AB.

XX Ahman I, Melander M, Vamling K;

XX WPI; 2000-160693/14.

XX N-PSDB; AAZ58018.

XX
 PT

Novel lectins used to produce transgenic Brassica plants which are resistant to insects

XX Example 2a; Fig 6; 51pp; English.

XX

CC The present sequence is that of a jack bean concanavalin A
 CC (Con A) mutein in which the Asn-152 residue of the native protein
 CC is replaced by Ser in order to disrupt a glycosylation signal. The
 CC present invention relates to: a transformed Brassica plant that is
 CC resistant to certain insect pests; an expression cassette
 CC containing DNA that codes for at least 1 lectin selected from Con A
 CC (see AAY58736), modified Con A and pea lectin (see AAY58738); transgenic
 CC plant cells containing at least 1 copy of the DNA; the present
 CC lectin derived from the jack bean Con A gene; a method of imparting
 CC resistance to insects selected from blossom beetles (pollen
 CC beetles) of the genus *Meligethes*, flea beetles of the genus
 CC *Phyllotreta*, and root flies of the genus *Delia*; and a method for
 CC protecting a plant against infestation by insects of these genera.
 XX

SQ Sequence 290 AA;

Query Match 40.8%; Score 653; DB 21; Length 290;
 Best Local Similarity 51.9%; Pred. No. 1e-52;

Matches 139; Conservative 43; Mismatches 66; Indels 20; Gaps 9;

QY 3 SLSEFNTKFDLQKDLIFQGDATS--TNNVLQTLKLDGNGPVGASVGRVLSAPFHLWEN 61

DB 36 ALHFMFNQFSKQDKLILQGDATTGTGDNLELTRVSSNGSPQGSVGRALFYAPVHIWES 95

QY 62 SMVSSFTETNLTIQISTP--HPYYAAGGFAFFLPHDTPVPNSWGKFLGLY--SNVFRN 117

DB 96 SAVVASFEATFTFLIKSPDSHP---ADGIAPFISNIDSSIPSGSTGRLLGLFPDANVIRN 152

QY 118 SPTSE--NQSFGDVNTDSRVVAVEFDTPPNANI--DPNVRHIGIDVNSIKSKETARWQNG 175

DB 153 STTIDFNAAY---NADT-IVAVELDTVPNTDIGDPSPYPHIGIDIKSVRSKKTAKNNWQNG 208

QY 176 KTATARISYNSASKKSTVTTFYFGMEVVALSHVDLHAELPEWVRVGLSASTGEEKQKNT 235

DB 209 KVGTAHIYNSVDKRLSAVSYPNADSATVSYDVLNDVLPFWVRVGLSASTGLYKETNT 268

QY 236 ILSWSTSSLLKNEVKEPKEDMYIANVV 263

DB 269 ILSWSTSKLSNEIPD-----IATVV 290

RESULT 8

AAR74765

ID AAR74765 standard; protein; 237 AA.

XX AAR74765;

XX 13-OCT-1995 (first entry)

XX Legume concanavalin A.

XX Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;
 XX amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;
 XX female hamster protein; 3-D structure; jellyroll topology;
 XX picornavirus; coat protein; pentameric structure; concanavaline A.

XX Legume sp..

XX WO9505394-A.

XX 23-FEB-1995.

XX 17-AUG-1994; 94WO-GB01802.

XX 17-AUG-1993; 93GB-0017120.

XX (BIRK-) BIRKBECK COLLEGE.

PN WO9911278-A1.
 XX 11-MAR-1999.
 XX 28-AUG-1998; 98WO-GB02612.
 XX 29-AUG-1997; 97GB-0018413.
 XX (ALIZ-) ALIZYME THERAPEUTICS LTD.
 XX Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
 XX WPI; 1999-228935/19.
 XX Use of Robinia pseudoacacia lectin in medicine
 XX Disclosure; Fig 2; 37pp; English.
 XX The invention relates to Robinia pseudoacacia lectins which can be used
 CC in medicine. The lectin is used for the control of mucosal cell
 CC proliferation, for the reduction and/or treatment of damage caused by
 CC cell damaging agents, especially in the treatment of cancer, and/or for
 CC the reduction and/or treatment of metabolic disorders. It is especially
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,
 CC particularly human cells and/or tissues (especially mucous cells
 CC including mucous membrane). Use of the lectin is effective and does not
 CC damage the gut as is the case with cytotoxic drugs and radiation. The
 CC present sequence represents a R. pseudoacacia lectin RPBAl. This lectin
 CC is composed of two subunits- polypeptide a and polypeptide b
 CC (AA06811-12).
 XX
 SQ Sequence 285 AA;
 Query Match 30.6%; Score 489.5; DB 20; Length 285;
 Best Local Similarity 43.6%; Pred. No. 2.2e-37;
 Matches 113; Conservative 37; Mismatches 86; Indels 23; Gaps 8;
 QY 3 SLSEFNFKFDLDKQLIFQGDATSTN-NVLQLTKLDGSGNPVGASVGRVLSAPFHLWEN 61
 Db 34 SLSEFPFKFAPNPQYLIFQRDALVTSGVLQLTNV-VNGVPPRRSIRGALYAAPFQIWDN 92
 QY 62 SMA-VSSFTETNLTIQISTPHPYAAADGFAFFLAPHDVTIPPNMSGKFLGLYSNVFRNSPT 120
 Db 93 TTGNVASFVTSFSEFIQAPNPATTAADGLAFAPVDT--QPGDLGGMLGIFD----- 143
 QY 121 SENQSGDVNTDSRVAVEDFTFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180
 Db 144 -----GSYNKSNQIVAVEFDFTSNIFHDPKGRHMGINVNSIVSKVTPVNNWNTNCEVANV 197
 QY 181 RISYNSAKSKSTVTTPYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236
 Db 198 FISYEASTKSLNASLVPSLETSTFIHAIVDVKVLPPEWVRVGFSAATTGIDTYVQTNV 257
 QY 237 ISWSTSSLK-NNEVKPEK 254
 Db 258 LSWSFESNLPGGNSVASVK 276
 RESULT 11
 AAY06812
 ID AAY06812 standard; protein; 285 AA.
 XX
 AC AAY06812;
 XX
 DT 23-JUN-1999 (first entry)
 XX
 DE R. pseudoacacia lectin RPBAl polypeptide subunit b.
 XX
 DE Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
 KW cancer; metabolic disorder; mucositis; cytotoxic.
 XX
 OS Robinia pseudoacacia.
 XX

PN WO9911278-A1.
 XX 11-MAR-1999.
 XX 28-AUG-1998; 98WO-GB02612.
 XX 29-AUG-1997; 97GB-0018413.
 XX (ALIZ-) ALIZYME THERAPEUTICS LTD.
 XX Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
 XX WPI; 1999-228935/19.
 XX Use of Robinia pseudoacacia lectin in medicine
 XX Disclosure; Fig 1; 37pp; English.
 XX The invention relates to Robinia pseudoacacia lectins which can be used
 CC in medicine. The lectin is used for the control of mucosal cell
 CC proliferation, for the reduction and/or treatment of damage caused by
 CC cell damaging agents, especially in the treatment of cancer, and/or for
 CC the reduction and/or treatment of metabolic disorders. It is especially
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,
 CC particularly human cells and/or tissues (especially mucous cells
 CC including mucous membrane). Use of the lectin is effective and does not
 CC damage the gut as is the case with cytotoxic drugs and radiation. The
 CC present sequence represents a polypeptide subunit of R. pseudoacacia
 CC lectin RPBAl.
 XX
 SQ Sequence 285 AA;
 Query Match 30.4%; Score 485.5; DB 20; Length 285;
 Best Local Similarity 44.8%; Pred. No. 5.1e-37;
 Matches 116; Conservative 39; Mismatches 81; Indels 23; Gaps 9;
 QY 3 SLSEFNFKFDLDKQLIFQGDATSTN-NVLQLTKLDGSGNPVGASVGRVLSAPFHLWEN 61
 Db 34 SLSEFPFKFHKHSDPLIFQSDALVTSGVLQLTNV--RVYDSIGRVLAAAPFQIWDN 91
 QY 62 SMA-VSSFTETNLTIQISTPHPYAAADGFAFFLAPHDVTIPPNMSGKFLGLYSNVFRNSPT 120
 Db 92 TTGNVASFVTSFSEFIQAPNPATTAADGLAFAPVDT--QPLKGGGLGLF----- 141
 QY 121 SENQSGDVNTDSRVAVEDFTFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180
 Db 142 -KDESY--NKSQIVAVEFDFTFPNANIDPNYRHIGIDVNSIQSVTRVWDWANGEVANV 197
 QY 181 RISYNSAKSKSTVTTPYPGME-VVALSHDVLHAELPEWVRVGLSASTG---EEKQNTI 236
 Db 198 FISYEASTKSLNASLVPSLEKSFILSAIVDLKKVLPPEWVRVGFATTGLSEYVQTNV 257
 QY 237 ISWSTSSLK-NNEVKPEK 254
 Db 258 LSWSFESNLPGGNSVASVK 276
 RESULT 12
 AAY06815
 ID AAY06815 standard; protein; 285 AA.
 XX
 AC AAY06815;
 XX
 DT 23-JUN-1999 (first entry)
 XX
 DE Amino acid sequence of lectin RPBAl from R. pseudoacacia seed.
 XX
 DE Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
 KW cancer; metabolic disorder; mucositis; cytotoxic.
 XX
 OS Robinia pseudoacacia.
 XX
 PN WO9911278-A1.

XX PD 11-MAR-1999.
 XX PF 28-AUG-1998; 98WO-GB02612.
 XX PR 29-AUG-1997; 97GB-0018413.
 XX PA (ALIZ-) ALIZYME THERAPEUTICS LTD.
 XX PI Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
 XX WPI; 1999-228935/19.
 XX DR Use of Robinia pseudoacacia lectin in medicine
 XX PT Disclosure; Fig 2; 37pp; English.
 XX PS
 XX CC The invention relates to Robinia pseudoacacia lectins which can be used
 CC in medicine. The lectin is used for the control of mucosal cell
 CC proliferation, for the reduction and/or treatment of damage caused by
 CC cell damaging agents, especially in the treatment of cancer, and/or for
 CC the reduction and/or treatment of metabolic disorders. It is especially
 CC useful for the treatment of mucositis in mammalian cells and/or tissues,
 CC particularly human cells and/or tissues (especially mucous cells
 CC including mucous membrane). Use of the lectin is effective and does not
 CC damage the gut as is the case with cytotoxic drugs and radiation. The
 CC present sequence represents a R. pseudoacacia lectin RPB11. This lectin
 CC is composed of the subunit- polypeptide c (AAY06813).
 XX CC
 XX SQ Sequence 285 AA;

Query Match 30.2%; Score 483.5; DB 20; Length 285;
 Best Local Similarity 44.0%; Pred. No. 7.9e-37;
 Matches 114; Conservative 37; Mismatches 85; Indels 23; Gaps 8;

QY 3 SLSEFTFDLDDKDLIFQGDATSTN-NVLQTLKDSGPNPGASVGRVLFSAPEHLWEN 61
 DB 34 SLSEFTFPAPNPQYLIFQDALVTSTGVQLTNV-VNGVPSKSLGRALYAPFQIWD 92
 QY 62 SMA-VSSFTNLTIQISTPHYPYAADGFAFFLAPHDTVIPNSWGKFLGLYSNVFRNSPT 120
 DB 93 TTGNVASEVTSFSTFIQAPNPATTADGLAFLAPVDV--QPLDLGGMGLGIFKN----- 143
 QY 121 SENQSGDVNTDSRVAVVEFDTPFNANIDPNYRHIGIDVNSIKSKETARWONGKATA 190
 DB 144 -----GYFNKSNQIVAVEFDTFNRHWDPTGRHMGINVSIVSVKTVPNWANGEVANV 197
 QY 181 RITSYNSASKKSTVTTFYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEK---OKNTI 236
 DB 198 FTSYASTKSLTASLYPSLETSFIHAIVDVKDVLPEWVRFGFSATTGIDTGYVQTNV 257
 QY 237 ISWFTSSLK-NNEVKEPK 254
 DB 258 LSWSPSNLPGGNSVASVK 276

RESULT 13
 AAY58738
 ID AAY58738 standard; Protein; 275 AA.
 XX AC AAY58738;
 XX DT 25-APR-2000 (first entry)
 XX DE Pea lectin.
 XX KW Lectin; pea; insecticide; transgenic plant; Brassica;
 XX KW insect resistance.
 XX OS Pisum sativum.
 XX PN WO20000123-A1.
 XX OS

PD 13-JAN-2000.
 XX PF 02-JUL-1999; 99WO-SE01209.
 XX PR 03-JUL-1998; 98SE-0002425.
 XX PA (PLAN-) PLANT SCI SVERIGE AB.
 XX PI Ahman I, Melander M, Vamling K;
 XX WPI; 2000-160693/14.
 XX DR N-PSDB; AAZ58019.
 XX PT Novel lectins used to produce transgenic Brassica plants which are
 XX resistant to insects
 XX PS Example 2b; Fig 9; 51pp; English.
 XX CC The present sequence is that of pea cv. Lincoln lectin. The
 CC invention relates to: a transformed Brassica plant that is
 CC resistant to certain insect pests; an expression cassette
 CC containing DNA that codes for at least 1 lectin selected from Con A
 CC (see AAY58736), modified Con A (see AAY58737) and pea lectin;
 CC transgenic plant cells containing at least 1 copy of the DNA; a new
 CC lectin derived from the Jack bean Con A gene; a method of imparting
 CC resistance to insects selected from blossom beetles (pollen
 CC beetles) of the genus Meligethes, flea beetles of the genus
 CC Phyllotreta, and root flies of the genus Delia; and a method for
 CC protecting a plant against infestation by insects of these genera.
 XX CC
 XX SQ Sequence 275 AA;

Query Match 29.5%; Score 471; DB 21; Length 275;
 Best Local Similarity 40.5%; Pred. No. 1.1e-35;
 Matches 107; Conservative 42; Mismatches 85; Indels 30; Gaps 6;

QY 2 QLSFNTKEDLDKDLIFQGDATSTNNVLQTLKDSGPNPGASVGRVLFSAPEHLWEN 61
 DB 32 ETTSFLTKRTPDOONLIFQGDGYTTREKLTAK-----AVKNTVGRALYSSPFIHWD 85
 QY 62 SMA-VSSFTNLTIQISTPHYPYAADGFAFFLAPHDTVIPNSWGKFLGLYSNVFRNSPT 120
 DB 86 ETGNVAVNFVTSFTEVINAPNSYNVADGETTFIAPVDV--KPQTGGGVLGVFNS----- 136
 QY 121 SENQSGDVNTDSRVAVVEFDTPFNANIDPNY--RHIGIDVNSIKSKETARWONGKATA 178
 DB 137 -----AEYDKTTQTAVAVEFDTFYNAWDPSNRDRHIGIDVNSIKSVNTKSKLQNGEEA 190
 QY 179 TARTSYNSASKKSTVTTFYP-----GMEVVALSHDVLHAELPEWVRVGLSASTGEKQ 232
 DB 191 NVVTAFAATNVLTVSLTYPNSLEENVTSTLSDVSLKDVVPEWVRFGFSATTGAEYA 250
 QY 233 KNTIISWTSSTSLKNNVEKPEK 256
 DB 251 AHEVLSWSFHSLSGTSSTSSKQAAD 274

RESULT 14
 AAR74766
 ID AAR74766 standard; protein; 228 AA.
 XX AC AAR74766;
 XX DT 13-OCT-1995 (first entry)
 XX DE Pea lectin, 2ltn.
 XX KW Pentraxin; legume; lectin; structural homology; serum amyloid P; SAP;
 XX KW amyloid; fibril; human; Limulus; C-reactive protein; CRP; pea;
 XX KW female hamster protein; 3-D structure; jellyroll topology;
 XX KW picornavirus; coat protein; pentameric structure; concanavaline A.
 XX OS Pisum sativum.

XX WO9505394-A.
 XX 23-FEB-1995.
 XX 17-AUG-1994; 94WO-GB01802.
 XX 17-AUG-1993; 93GB-0017120.
 XX (BIRK-) BIRKBECK COLLEGE.
 XX (ROYA-) ROYAL POSTGRAD MED SCHOOL.
 PI Blundell TL, Pepys MB;
 DR WPI; 1995-098720/13.
 XX New cpds. Inhibiting binding of serum amyloid P to amyloid
 PT fibrils - produced by computer assisted molecular design, useful
 PT for preventing, treating or diagnosing amyloidosis, e.g.
 PT Alzheimer's disease
 XX
 PS Claim 7; Fig 4; 72pp; English.
 XX
 CC The sequences given in AAR74763-70 represent various pentraxins and
 CC legume lectins. All these proteins show structural homology and may
 CC be used in the molecular design of a molecule for the inhibition of
 CC serum amyloid P (SAP) binding to amyloid fibrils. The similarities
 CC in the amino acid sequences of SAP, human and Limulus C-reactive
 CC protein (CRP) and female hamster protein suggests that they may have
 CC similar 3-D structures. The jellyroll topology of the pentraxins is
 CC reminiscent of the picornavirus coat proteins which also have
 CC pentameric structures. However, pentaxins resemble more closely
 CC legume lectins such as concanavaline A and pea lectin. Alignment
 CC of sequences on the basis of topologically equivalent features of the
 CC three dimensional structures shows that helices occupy different
 CC positions in the pentraxins and legume lectins and that the amino acid
 CC sequences of the two families have identities of only approx. 11%. The
 CC two main helices in SAP occur before and after strand L, whereas the
 CC helices in the legume lectins occur at the C-terminus of strand J.
 CC There is a long insertion between the end of the helix after strand D
 CC and the beginning of strand E in the lectins relative to pentraxins.
 CC Strands G, H and I together with the type IV beta-hairpin between H and
 CC I are identical in both SAP and pea lectin. The so-called pentraxin
 CC octapeptide signature sequence, HXCXS/TWXS, is in this region so it is
 CC not conserved in the legume lectins.
 XX
 SQ Sequence 228 AA;
 Query Match 26.0%; Score 416; DB 16; Length 228;
 Best Local Similarity 41.3%; Pred. No. 1.2e-30;
 Matches 92; Conservative 38; Mismatches 69; Indels 24; Gaps 5;
 QY 26 STNNVQLTKLDSGGNPVGASVGRVLSAPFHLWNSMA-VSSPETNLTIQISTPHPYA 84
 Db 2 TTEKELTLTK-----AKVTVGRALYSFIIHWDRTGNVAVFTFTFVLNAPNSVNV 55
 QY 85 ADGFAFLAPHDVIPPNSWGKFLGLYSNVFNPSPTSENQSGDVNTDSRVVAVEEDTFP 144
 Db 56 ADGFTFFIAPVDV--KPQTGGVGLVFN-----AEYDKTQTVVAVEEDTFY 100
 QY 145 NANIDPNY--RHIGIDVNSKSKETARWQNGKTATARYNSASKSVITTFYPGMEV 202
 Db 101 NAAMDVSNRDRHIGIDVNSKSKNTKSWKLONGEEANVIAFNATNVLTSLTYPNVT 160
 QY 203 VALSHVDVLHAEFLPEWVRVGLSASTGEEKQNTIISNFTSSL 245
 Db 161 YTLSDVVVSLKDVPEWVRIGFSAITGAEYAHEVLSNFSHSEL 203
 RESULT 15
 AAY27265
 ID AAY27265 standard; protein; 262 AA.
 XX

AC AAY27265;
 XX 22-OCT-1999 (first entry)
 XX B. purpurea native lectin sequence.
 XX Bauhinia purpurea lectin; lectin; BPL; larvicide; insect control;
 KW European corn borer; Ostrinia nubilalis; crop protection;
 KW insect larvae infestation.
 OS Bauhinia purpurea.
 XX US5945589-A.
 XX 31-AUG-1999.
 XX 24-MAR-1993; 93US-0038761.
 XX 24-MAR-1993; 93US-0038761.
 PR 24-JUL-1992; 92US-0921179.
 XX (PION-) PIONEER HI-BRED INT INC.
 XX Balasubramaniam NK, Rao AG;
 XX WPI; 1999-508204/42.
 XX Derivatives of Bauhinia purpurea lectins useful as larvicides to
 PT control insects (e.g. European corn borer) in field crops such as
 PT corn, wheat and rapeseed
 XX Example 5; Columns 11-14; 8pp; English.
 PS The invention describes compounds derived from Bauhinia purpurea lectin
 CC ((BPL), a 262 amino acid peptide comprising 7 lysine residues - the
 CC present sequence) by the replacement of one or more lysine residues with
 CC other amino acid residues which either preserve the positive charge or
 CC provide a neutral residue at the point of substitution. The compounds are
 CC useful as larvicides against insects such as the European corn borer
 CC (Ostrinia nubilalis). The compounds may be used to protect commercial
 CC field crops from insect larvae infestation, especially by the European
 CC corn borer. The BPL-derivatives are effective against insects which
 CC may be resistant to the naturally occurring form of BPL. Crops which may
 CC be protected using the BPL derivatives are corn/maize (Zea mays), rye,
 CC barley, wheat (Triticum aestivum), sorghum (Sorghum bicolor), oats,
 CC millet, rice, triticale, sunflower, alfalfa, rapeseed and soybean), fiber
 CC crops (such as cotton), fruit crops (such as melons) and vegetable crops
 CC (such as onion, pepper, tomato, cucumber, squash, carrot, crucifer (e.g.
 CC cabbage, broccoli and cauliflower), eggplant, spinach, potato and
 CC lettuce). The compounds are useful for protecting species from the genera
 CC Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna,
 CC Citrus, Linum, Geranium, Manicot, Daucus, Arabidopsis, Brassica,
 CC Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon,
 CC Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus,
 CC Lactuca, Bromus, Asparagus, Antirrhinum, Hemerocallis, Nemesis,
 CC Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis,
 CC Cucumis, Browallia, Glycine, Lolium, Triticum and Datura. The present
 CC sequence represents the B. purpurea native lectin sequence.

SQ Sequence 262 AA;

Query Match 24.0%; Score 383; DB 20; Length 262;
 Best Local Similarity 40.3%; Pred. No. 1.8e-27;
 Matches 96; Conservative 39; Mismatches 85; Indels 18; Gaps 7;
 QY 17 DLIFOGDATSTNNVQLTKLDSGGNPVGASVGRVLSAPFHLWNSMAVSFETNLTIQI 76
 Db 23 ELIFLGNATYTPGALTRIGEDGIPKSNAGASVSRPVFLWDSTGHVASFTSPSFIV 82
 QY 77 -STPHPYNAADGFAFLAPHDVIPPNSWGKFLGLYSNVFNPSPTSENQSGDVNTDSRV 135
 Db 83 RSIDVPHITADGFAFLAPVDSSV--KDYGGCLGLFRYKTATDP-SKNQ-----V 129

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 ; Search time 28.7491 Seconds
(without alignments)
2171.628 Million cell updates/sec

Title: US-09-476-485A-6
Perfect score: 1599
Sequence: 1 AQSLSFNFTKFDLDQKDLIF.....LNNHKYVRCSTCMLFMKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rviro.*
- 16: sp_bacterioph.*
- 17: sp_archaeap.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO	Score	Query Match %	Length	ID	Description
1	1464	91.6	279	10 Q9M7M4	Q9m7m4 phaseolus v
2	804	50.3	272	10 Q9ZTA9	Q9zta9 dolichos la
3	658	41.2	290	10 Q947H0	Q947h0 canavalia e
4	656	41.0	290	10 Q04672	O04672 canavalia b
5	589.5	36.9	266	10 P93536	P93536 sophora jap
6	582.5	36.4	284	10 Q9FYU9	Q9fyu9 sophora fla
7	579.5	36.2	293	10 P93537	P93537 sophora jap
8	513	32.1	286	10 P93248	P93248 maackia amu
9	510	31.9	282	10 Q93247	Q93247 maackia amu
10	504	31.5	254	10 Q43376	Q43376 arachis hyp
11	498.5	31.2	285	10 Q9ZWP6	Q9zwp6 robinia pse
12	489	30.6	280	10 Q43374	Q43374 arachis hyp
13	481	30.1	254	10 Q43377	Q43377 arachis hyp
14	477	29.8	279	10 Q49899	Q49899 medicago sa
15	471	29.5	251	10 Q93X48	Q93x48 lens ervoid
16	467	29.2	251	10 Q8RW33	Q8rw33 lathyrus sa

17	465	29.1	251	10 Q93X42	Q93x42 lens nigric
18	465	29.1	251	10 Q93WH6	Q93wh6 lens culina
19	464	29.0	251	10 Q93X41	Q93x41 lens odemen
20	464	29.0	251	10 Q8W4Y4	Q8w4y4 lens lamott
21	462	28.9	251	10 Q8VXF2	Q8vxf2 lens culina
22	461	28.8	251	10 Q93X50	Q93x50 lens culina
23	461	28.8	251	10 Q93X49	Q93x49 lens culina
24	457.5	28.6	275	10 Q8RW23	Q8rw23 phaseolus c
25	456	28.5	247	10 Q8W1R6	Q8w1r6 griffonia s
26	456	28.5	258	10 Q9FVF8	Q9fvf8 ulex europe
27	456	28.5	275	10 Q8RVY1	Q8rvy1 phaseolus v
28	455.5	28.5	251	10 Q8W1R7	Q8w1r7 griffonia s
29	452.5	28.3	256	10 P93246	P93246 maackia amu
30	451.5	28.2	273	10 Q8RVY4	Q8rvy4 phaseolus c
31	446.5	27.9	258	10 Q41263	Q41263 griffonia s
32	444.5	27.8	276	10 Q40750	Q40750 phaseolus a
33	439.5	27.5	275	10 Q8RVH1	Q8rvh1 phaseolus v
34	438.5	27.4	275	10 Q8RVH3	Q8rvh3 phaseolus v
35	437.5	27.4	275	10 Q8RVY0	Q8rvy0 phaseolus v
36	437.5	27.4	275	10 Q8RVX6	Q8rvx6 phaseolus v
37	435.5	27.2	273	10 Q8RVH2	Q8rvh2 phaseolus v
38	435.5	27.2	278	10 Q9LED8	Q9led8 phaseolus l
39	433.5	27.1	273	10 Q8RVX9	Q8rvx9 phaseolus v
40	430.5	26.9	273	10 Q42411	Q42411 medicago sa
41	430.5	26.9	278	10 Q9LED9	Q9led9 phaseolus l
42	430.5	26.9	285	10 Q9ZWP5	Q9zwp5 robinia pse
43	430	26.9	274	10 Q43628	Q43628 phaseolus v
44	429	26.8	259	10 Q8RVW9	Q8rvw9 vicia faba
45	428.5	26.8	268	10 P93458	P93458 phaseolus l

ALIGNMENTS

RESULT 1

Q9M7M4 ID Q9M7M4 PRELIMINARY; PRT; 279 AA.
AC Q9M7M4;
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE Mannose lectin FRIL (Fragment).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
RA Feldman M.;
RT "A new lectin in red kidney bean called PvFRIL stimulates
RT proliferation of NIH3T3 cells expressing the Flt3 receptor."
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121458; AAF28739.1;
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
FT NON-TER
SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;

Query Match 91.6%; Score 1464; DB 10; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.6e-106;
Matches 279; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 AQSLSFNFTKFDLDQKDLIFQGDATSTNNVLQTLKLDGSGNPVGASVGRVLSAPFHLWE 60

```

1 AQLSFTNFKFDLQKOLIFQGDATSTNNVLQTLKLDGGNPVGASVGRVLSAPHLWE 60
61 NSMAVSSSETNLTIQISTPHTPHYYAAGDAFFLAPHDVTIPPNSWGKFLGLYSNVFRNSPT 120
61 NSMAVSSSETNLTIQISTPHTPHYYAAGDAFFLAPHDVTIPPNSWGKFLGLYSNVFRNSPT 120
121 SENOSFGDVNTDSRVVAVFEFTFPNANTDPNRYRHIGIDVNSIKSKETARWONGKTATA 180
121 SENOSFGDVNTDSRVVAVFEFTFPNANTDPNRYRHIGIDVNSIKSKETARWONGKTATA 180
181 RISYNSASKSKSTVTTFYPCMEVVVALSHDVLHAELEPWWVRVGLSASTGEEKOKNTIISWS 240
181 RISYNSASKSKSTVTTFYPCMEVVVALSHDVLHAELEPWWVRVGLSASTGEEKOKNTIISWS 240
241 FTSSLKKNNEVKPEKEDMYIANVVRYSYTWINDVLSYLSNK 279
241 FTSSLKKNNEVKPEKEDMYIANVVRYSYTWINDVLSYLSNK 279

RESULT 2
Q9ZTA9 PRELIMINARY; PRT: 272 AA.
AC Q9ZTA9;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Mannose lectin.
GN DOLICHOS
OS Dolichos lab lab (Field bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=35936;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RX COLLINS-9910944; PubMed=9892687;
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;
RT "cDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves
RT hematoietic progenitors in suspension culture.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).
DR EMBL; AF067417; AAD10734.1; -.
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 50.3%; Score 804; DB 10; Length 272;
Best Local Similarity 60.9%; Pred. No. 8.3e-55;
Matches 162; Conservative 39; Mismatches 59; Indels 6; Gaps 4;

QY 1 AQLSFTNFKFDLQKOLIFQGDATSTNNVLQTLKLDGGNPVGASVGRVLSAPHLWE 60
9 AQLSFTNFKFDLQKOLIFQGDATSTNNVLQTLKLDGGNPVGASVGRVLSAPHLWE 68
61 NSMAVSSSETNLTIQISTPHTPHYYAAGDAFFLAPHDVTIPPNSWGKFLGLYSNVFRNS 118
69 DSAVLTSTDTIINFESTPTYSRIADGLAFTAPPDSVI--SYHGGFLGLFPNANTLNNS 126
119 PTSENOSFGDVNTDSRVVAVFEFTFPNANTDPNRYRHIGIDVNSIKSKETARWONGKT 177
127 STSENQTTTKA-ASSNVAVFEFTYLPDYPDNPYIHIGIDVNSIRSKYTAQWQNGKI 185
178 ATARISYNSASKSKSTVTTFYPCMEVVVALSHDVLHAELEPWWVRVGLSASTGEEKQNTII 237
178 ATARISYNSASKSKSTVTTFYPCMEVVVALSHDVLHAELEPWWVRVGLSASTGEEKQNTII 237

```

```

186 ATAHISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPEWVRVGLSASTGQDKRNTVH 245
238 SWSFTSSLKKNNEVKPEKEDMYIANVV 263
246 SWSFTSSLTWTNVAKKENENKIYITRGV 271

RESULT 3
Q947H0 PRELIMINARY; PRT: 290 AA.
AC Q947H0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Concanavalin A.
GN CANAVALIA
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. U-02;
RA Ramis C., Gomord V.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308777; AAL09432.1; -.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
SQ SEQUENCE 290 AA; 31480 MW; 0F2F7DBBCF547E42 CRC64;

Query Match 41.2%; Score 658; DB 10; Length 290;
Best Local Similarity 52.2%; Pred. No. 2.1e-43;
Matches 140; Conservative 42; Mismatches 66; Indels 20; Gaps 9;

QY 3 SLSENFNFKFDLQKOLIFQGDATSTNNVLQTLKLDGGNPVGASVGRVLSAPHLWEN 61
36 ALHFMFNQFSKDKDLILQGDATTTGDNLELTRVSSNGSQSGSVGRALFYAPVHIES 95
62 SMAVSSSETNLTIQISTP--HPYYAAGDAFFLAPHDVTIPPNSWGKFLGLY--SNVFRN 117
96 SAVVASFEATFTFLIKSPDSHP---ADGIAFFISNIDSIDSPSGTGRLLGLFPDANVRN 152
118 SPTSE-NQSGFDVNTDSRVVAVFEFTFPNANTDPNRYRHIGIDVNSIKSKETARWONG 175
153 STTIDFNAAY---NADT-IVAVELDTYPTDIDGPPSYPHIGIDIKSVRSKKTAKWQMNG 208
176 KTATARISYNSASKSKSTVTTFYPCMEVVVALSHDVLHAELEPWWVRVGLSASTGEEKQNT 235
209 KVGFAHIYNSVDKRLSAVSVYPNADSATVSYDVLDMVLPEWVRVGLSASTGLYKETNT 268
236 IISWSFTSSLKKNNEVKPEKEDMYIANVV 263
269 ILSWSFTSKLSNEIPD-----IATVV 290

RESULT 4
O04672 PRELIMINARY; PRT: 290 AA.
AC O04672;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Lectin (Fragment).
OS Canavalia brasiliensis (Brazilian jack bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```


OC Spermatozoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.
 OX NCBI_TaxID=37501;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-48.
 RC TISSUE=BARK;
 RX MEDLINE=97390228; PubMed=9249142;
 RA van Damme E.J.M., Van Leuven F., Peumans W.J.;
 RT "Isolation, characterization and molecular cloning of the bark lectins
 RL Glycoconj. J. 14:449-456(1997).
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
 CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 DR EMBL; U65010; AAB39934.1; -;
 DR HSP; P19588; ILUL.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR InterPro; IPR000531; TonB_boxC.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 DR LECTIN; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
 FT NON_TER 1
 FT SIGNAL <1 28
 FT CHAIN 29 286 BARK LEUCOAGGLUTININ II.
 FT CARBOHYD 27 27 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 286 AA; 31208 MW; 0F2E27617A0F6D00 CRC64;
 Query Match 32.1%; Score 513; DB 10; Length 286;
 Best Local Similarity 43.0%; Pred. No. 4e-32;
 Matches 116; Conservative 51; Mismatches 77; Indels 26; Gaps 9;
 QY 1 AQSLSFETKFDLDKDLIFOGDAT-STNNVLQTLKDSGNGPVGASVGRVLFSAFFHLW 59
 Db 29 SDELSFTINFPNEADLLFOGEASVSTGVQLTRVEN-GQPQKYSVGRALYAAPVRW 87
 QY 60 ENSM-AVSSFETNLTIQISTPHPYAAGDAFFLAHDVTIPPNWSGKFLGLYSNVFRNS 118
 Db 88 DNTTGSVASFSTFTFVVKAPNPITSDGLAFLAPDSQIPSRVSKYLGLEN--SNS 145
 QY 119 PTSENOQSGDVNTDSRVVAVEFTFPNANI---VVALSHDVLHAELPEWVRVGLSASTG 175
 Db 146 DSS-----NQIVAVEFTFYGHSDPDPNRYRHIGIDVNGIESIKTVQMDWING 194
 QY 176 KTATARISNSAKSKSTVTTFYPMGE---VVALSHDVLHAELPEWVRVGLSASTG 232
 Db 195 GVAFATITVLPKNTLIASLVYFSNQTSTFIVAAS--VDLKEILPEWVRVGFSAATGYPTQ 252
 QY 233 KNT--IISWFTSSLNKNEVKPEKEDMYIA 260
 Db 253 VETHDVLWSWFTSTLEANSDAATENNVAIA 282
 RESULT 9
 P93247 PRELIMINARY; PRT; 282 AA.
 AC P93247
 DT 01-MAY-1997 (Tremblrel. 03, Created)
 DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Bark LEUCOAGGLUTININ I precursor (MALBI) (Fragment).
 GN LECWALBT
 OS Maackia amurensis.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatozoa; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.
 OX NCBI_TaxID=37501;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BARK;
 RX MEDLINE=97390228; PubMed=9249142;
 RA van Damme E.J.M., Van Leuven F., Peumans W.J.;
 RT "Isolation, characterization and molecular cloning of the bark lectins
 RL Glycoconj. J. 14:449-456(1997).
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
 CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 DR EMBL; U65009; AAB39933.1; -;
 DR HSP; P19588; ILUL.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR InterPro; IPR000531; TonB_boxC.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 DR LECTIN; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
 FT NON_TER 1
 FT SIGNAL <1 24 BY SIMILARITY.
 FT CHAIN 25 282 BARK LEUCOAGGLUTININ I.
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 282 AA; 30879 MW; 45B3F714EID8957A CRC64;
 Query Match 31.9%; Score 510; DB 10; Length 282;
 Best Local Similarity 42.9%; Pred. No. 6.7e-32;
 Matches 117; Conservative 52; Mismatches 72; Indels 32; Gaps 10;
 QY 1 AQSLSFETKFDLDKDLIFOGDAT-STNNVLQTLKDSGNGPVGASVGRVLFSAFFHLW 59
 Db 25 SDELSFTINFPNEADLLFOGEASVSTGVQLTRVEN-GQPQKYSVGRALYAAPVRW 83
 QY 60 ENSM-AVSSFETNLTIQISTPHPYAAGDAFFLAHDVTIPPNWSGKFLGLYSNVFRNS 118
 Db 84 DNTTGSVASFSTFTFVVKAPNPITSDGLAFLAPDSQIPSRVSKYLGLEN----- 138
 QY 119 PTSENOQSGDVNTDSRVVAVEFTFPNANI---VVALSHDVLHAELPEWVRVGLSASTG 172
 Db 139 -TSSDSS-----NQIVAVEFTFYGHSDPDPNRYRHIGIDVNGIDSIKTQVQMDW 187
 QY 173 QNGKTATARISNSAKSKSTVTTFYPMGE---VVALSHDVLHAELPEWVRVGLSASTG 228
 Db 188 INGVAFAITVLPKNTLIASLVYFSNQTSTFIVAAS--VDLKEILPEWVRVGFSAATGY 245
 QY 229 -EEKQNTIISWFTSSLNKNEVKPEKEDMYIA 260
 Db 246 PTEVETHDVLWSWFTSTLEANSDAATENNVAIA 278
 RESULT 10
 Q43376

ID Q43376 PRELIMINARY; PRT; 254 AA.
AC Q43376;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Mannose/glucose-binding lectin precursor (fragment).
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OC NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. SELLIE; TISSUE=SEED;
RC Law I.J.;
RT "Cloning and expression of cDNA for mannose-binding lectin from
peanut.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22472; AAA74575.1; -.
DR HSSP; P02867; 2BOP.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT NON_TER 1
FT SIGNAL <1
SQ SEQUENCE 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;

Query Match 31.5%; Score 504; DB 10; Length 254;
Best Local Similarity 45.4%; Pred. No. 1.7e-31;
Matches 113; Conservative 46; Mismatches 72; Indels 18; Gaps 7;

QY 3 SLSPFTKFDL-DKDLIPQGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFHLWE 60
DB 3 SLSPFTKFDL-DKDLIPQGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFHLWE 62
QY 61 NSM-AVSSFETNLTIQISTPHPIYAADGFAFFLAPHDVTIPPNWSGKFLGLYSNVFNPS 119
DB 63 KSNRLTNFQAQSFVSKSPND-IGADGIAFFAIPDQIPKNSAGGTGLF-----DPQ 116
QY 120 TSNQSGDVNTDSRVVAVFEFDTF---PNANIDPNYRHIGIDVNSIKSKETARWQNGK 176
DB 117 TAQNP-----ANQVAVEFDTFYQDSNGWDPNTQHIGIDVNSIKSAATTWERRNGQ 170
QY 177 TATARISYNSAKSKSTVTTPYPMGVVALSHDVLHAELPEWVRVGLSASTGEEKNTI 236
DB 171 TLNVLYTDANSKNLQVTSYPDGQRYQVSVVVDLRDHLPEWGRVGFSSSGQYQSHEL 230
QY 237 ISWSFTSSL 245
DB 231 QWSFTSNL 239

RESULT 11
Q92WP6 PRELIMINARY; PRT; 285 AA.
ID Q92WP6
AC Q92WP6;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Lactin.
OS Robinia pseudoacacia (Black locust).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
OC NCBI_TaxID=45938;
RN [1]

RP SEQUENCE FROM N.A.
RA Yoshida K.; Tazaki K.;
RT "Expression patterns of the genes that encode lectin or lectin-related
polypeptides in Robinia pseudoacacia".
RL Aust. J. Plant Physiol. 26:495-502(1999).
DR EMBL; AB012633; BAA36414.1; -.
DR HSSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
SQ SEQUENCE 285 AA; 30940 MW; 5B5A42C8B9579922 CRC64;

Query Match 31.2%; Score 498.5; DB 10; Length 285;
Best Local Similarity 44.8%; Pred. No. 5.3e-31;
Matches 116; Conservative 36; Mismatches 84; Indels 23; Gaps 8;

QY 3 SLSPFTKFDL-DKDLIPQGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFHLWE 61
DB 34 SLSPFTKFDL-DKDLIPQGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFHLWE 92
QY 62 SMA-VSSFETNLTIQISTPHPIYAADGFAFFLAPHDVTIPPNWSGKFLGLYSNVFNPSPT 120
DB 93 TTNVASEVTSFSTIQAPNPATADGLAFLAPVDT---QPLDLGMLGIFKN-----143
QY 121 SENQSGDVNTDSRVVAVFEFDTFPNANIDPNYRHIGIDVNSIKSKETARWQNGKTATA 180
DB 144 -----GYFNKSNQIIVAVEFDTFNRHNDPTGRHLGINVNSIKSVRTVPWNWTNGEVANV 197
QY 181 RSYNSAKSKSTVTTPYPMGVVALSHD-VLHAELPEWVRVGLSASTGEEK---QKNTI 236
DB 198 FTSYEASTKSLTASLVYPSLETSTFIHAIYVDVKDVLPEWVRVGLSASTGEEK---QKNTI 257
QY 237 ISWSFTSSLK-NNEVKEPK 254
DB 258 LSWSPFESNLPGGNSVASVK 276

RESULT 12
Q43374 PRELIMINARY; PRT; 280 AA.
ID Q43374
AC Q43374;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mannose/glucose-binding lectin precursor.
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OC NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-CV. SELLIE; TISSUE=MODULE;
RC Law I.J.;
RT "Cloning and expression of cDNA for mannose-binding lectin from
peanut.";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U22469; AAA74572.1; -.
DR HSSP; P02866; 1DQ2.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
RN [1]

KW Lectin; Signal.
 FT CHAIN 1 26
 SQ SEQUENCE 27 280 MANNOSYLGLUCOSE-BINDING LECTIN.
 SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;

Query Match
 Best Local Similarity 30.6%; Score 489; DB 10; Length 280;
 Matches 111; Conservative 45; Mismatches 75; Indels 18; Gaps 7;

QY 3 SLSFNFTKFDL-DOKDLIFOGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFFHLWE 60
 DB 29 SLSFSYNNEFQDDERNLILQGDATFSAGKGLTKVDNDGTPAKSTVGRVLSHTQVRLWE 88
 QY 61 NSM-AVSSFETNLTIQISTPHPYAADGFAFFLAHDVTIPPNWSWKGKFLGLYSNVRNRP 119
 DB 89 KSTNRLTNFOAQSFVKSIPID-NGADGIAFFIAAPDSEIPKNSAGGTGLGF-----DPS 142
 QY 120 TSENQSGDVNTDSRVVAVEDTF---PNANIDPNYRHIGIDVNSIKSKETARWQNGK 176
 DB 143 TAQNP-----ANQVLAVEEDTFYAQDSNGWDPNYQHIGIDVNSIKSAATTKWERRNGQ 196
 QY 177 TATARISVNSASKSTVTTFYPGMVVALSHDVLHAEPEWVRVGLSASTGEKQKNTI 236
 DB 197 TLNVLSYDANSKNLQVTASYPDQORYQVSYNVDRLDYLPEWGRVGFSAASQGYQSHL 256
 QY 237 ISWSFTSSL 245
 DB 257 QSWSFSTL 265

RESULT 13
 Q43377 PRELIMINARY; PRT; 254 AA.

ID Q43377
 AC Q43377
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)
 DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Manno9e/glucose-binding lectin precursor (Fragment).
 GN LEC.

OS Arachis hypogaea (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC Arachis.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SELLIE; TISSUE=SEED;
 RA Law I.J.;
 RT "Cloning and expression of cDNA for mannose/glucose-binding lectin
 from peanut.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U22473; AAA74576.1; -.
 DR HSSP; P02867; 2BQP.
 DR InterPro; IPR000985; Lectin_legA.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 KW LECTIN; Signal.
 FT SIGNAL
 SQ SEQUENCE 254 AA; 28389 MW; 761882817DE7DFB9 CRC64;

Query Match
 Best Local Similarity 30.1%; Score 481; DB 10; Length 254;
 Matches 110; Conservative 45; Mismatches 76; Indels 18; Gaps 7;

QY 3 SLSFNFTKFDL-DOKDLIFOGDAT-STNNVLQTLKLDGSGNPVGASVGRVLFSAFFHLWE 60
 DB 3 SLSFSYNNEFQDDERNLILQGDATFSAGKGLTKVDNDGTPAKSTVGRVLSHTQVRLWE 62

QY 61 NSM-AVSSFETNLTIQISTPHPYAADGFAFFLAHDVTIPPNWSWKGKFLGLYSNVRNRP 119
 DB 63 KSTNRLTNFOAQSFVKSIPID-NGADGIAFFIAAPDSEIPKNSAGGTGLGF-----DPO 116
 QY 120 TSENQSGDVNTDSRVVAVEDTF---PNANIDPNYRHIGIDVNSIKSKETARWQNGK 176
 DB 117 TAQNP-----ANQVLAVEEDTFYAQDSNGWDPNYQHIGIDVNSIKSAATTKWERRNGQ 170
 QY 177 TATARISVNSASKSTVTTFYPGMVVALSHDVLHAEPEWVRVGLSASTGEKQKNTI 236
 DB 171 TLNVLSYDANSKNLQVTASYPDQORYQVSYNVDRLDYLPEWGRVGFSAASQGYQSHL 230
 QY 237 ISWSFTSSL 245
 DB 231 QSWSFSTL 239

RESULT 14
 O49899 PRELIMINARY; PRT; 279 AA.

ID O49899
 AC O49899
 DT 01-JUN-1998 (T-EMBLrel. 06, Created)
 DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
 DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
 DE Lectin precursor.
 GN LEC2.

OS Medicago sativa (Alfalfa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 OX NCBI_TaxID=3879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CHIEF;
 RA Brill L.M.; Pieterneel V.R.;
 RT "Legume seed lectin genes: sequence of Mslec2 from Alfalfa, Alfalfa
 and Sweet Clover southern blot analysis, and stable transformation of
 Alfalfa with antisense-lectin constructs.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Y16754; CAA76366.1; -.
 DR HSSP; P04122; LLOE.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 KW Signal.
 FT SIGNAL
 FT CHAIN
 SQ SEQUENCE 279 AA; 30551 MW; 1315F022BABDA360 CRC64;

Query Match
 Best Local Similarity 29.8%; Score 477; DB 10; Length 279;
 Matches 111; Conservative 36; Mismatches 78; Indels 30; Gaps 6;

QY 1 AQSLSFNFTKFDL-DOKDLIFOGDA-TSTNNVLQTLKLDGSGNPVGASVGRVLFSAFFHLW 59
 DB 27 AETTSFSTKFPDQKNLIFQGDATTTSTGKLTK-----AVKNSIGRALYSAPTHW 80
 QY 60 ENSM-AVSSFETNLTIQISTPHPYAADGFAFFLAHDVTIPPNWSWKGKFLGLY-SNVRN 117
 DB 81 DSKTGSVANFETFTTITAPNTYINVADGLAFIADITOPKNSQGGYILGDFDSTYQE 140
 QY 118 SPTSENQSGDVNTDSRVVAVEDTFPNANIDP-----NYRHIGIDVNSIKSKETARW 170
 DB 141 S-----IQTVAVEIDTFYVNDWDPKPGNISSTGRHIGINVSISITVPW 186
 QY 171 EWQNGKTATARISVNSASKSTVTTFYPMVEVVALSHDVLHAEPEWVRVGLSASTGEE 230

Wed Feb 26 17:04:11 2003

Db 187 SLVNNKANVVGATNVLSDVEYPLVRHYTLSHVVPLKDVVPEWVRIGFSAATGDE 246
Qy 231 KOKNTIISWSTSSL 245
Db 247 YAEHDFSWFSDSKL 261

RESULT 15

Q93X48 PRELIMINARY; PRT; 251 AA.
AC Q93X48;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Lectin (Fragment).
GN LECTIN.
OS Lens ervoides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Lens.
OX NCBI_taxID=41257;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Galasso I.;
RT "Lectin genes in Lens."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ318220; CAC42126.1; -
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 251
SQ SEQUENCE 251 AA; 27963 MW; FE85E89765AF7DFD CRC64;

Query Match 29.5%; Score 471; DB 10; Length 251;
Best Local Similarity 42.6%; Pred. No. 6.2e-29;
Matches 106; Conservative 38; Mismatches 75; Indels 30; Gaps 6;
Qy 2 QSLSFNFTKFDLDOKDLIFQGDATSTNNVLQLKLDGSGNPVGASVGRVLFSAFPHLWEN 61
Db 23 ETTSSITKFPDQNLIFQGDGYTKKLTGK-----AVKNTVGRALYSTPIHIWDR 76
Qy 62 SMA-VSSFTNLTQISTHPHYAAGCAFFLAPHDVIPPNSWCKELGLYSNVFRNSPT 120
Db 77 DTGNVANFVTSFTFVINAPNSVNVADGTFPIAPVDT--KPTGGGYLGVFNS----- 127
Qy 121 SENQSGDVNTDSRVAVEDTFPNANIDPNY--RHIGIDVNSIKSKETARWQNGKTA 178
Db 128 -----KDYKTSQTVAVEFTFYNAAWDESNKDRHIGIDVNSIKSVKTSWNQNGERA 181
Qy 179 TARISNSAKSKSTVTTFYP-----GMEYVALSHVDLHAELEPEWVRVGLSASTGEKQ 232
Db 182 NVVIAENATNVLTVLTYPNSLEENVTSTLNEVVPMKDVLPWVRIGFSATTGAEEA 241
Qy 233 KNTIISWF 241
Db 242 AHEVLSWF 250

Search completed: February 26, 2003, 16:54:26
Job time : 29.7491 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:41:33 ; Search time 8.70037 Seconds
(without alignments)
1444.458 Million cell updates/sec

Title: US-09-476-485A-6

Perfect score: 1599

Sequence: 1 AQSLSFNFTKFDLDQKDLIF.....LNNHXYVRCTCMLFMKKK 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match %	Length	ID	Description
1	658	41.2	290	1	CONA_CANGL	P14894 canavalia g
2	656	41.0	290	1	CONA_CANEN	P02866 canavalia e
3	650	40.7	237	1	LECA_DOLLA	P38662 dolichos la
4	629.5	39.4	240	1	LEC_BOWMI	P42088 bowringia m
5	621	38.8	290	1	LEC2_GIALU	Q39529 cladrastis
6	604.5	37.8	293	1	LEC1_GIALU	Q39528 cladrastis
7	501.5	31.4	290	1	LECR_GIALU	Q39527 cladrastis
8	497	31.1	240	1	LECS_VATMA	P81371 vatireia ma
9	495	31.0	286	1	LEC2_ROBPS	Q42372 robinia pse
10	493.5	30.9	285	1	LCB1_ROBPS	Q41159 robinia pse
11	493	30.8	270	1	LECB_SOPJA	P93538 sophora jap
12	489.5	30.6	285	1	LCG1_ROBPS	Q41162 robinia pse
13	483.5	30.2	285	1	LCG2_ROBPS	Q41161 robinia pse
14	480	30.0	290	1	LECS_SOPJA	P93535 sophora jap
15	478	29.9	280	1	LEC2_WEDTR	Q01807 medicago tr
16	471	29.5	275	1	LEC_PEA	P02867 pisum sativ
17	470	29.4	248	1	LEC2_CYTSC	P29257 cytisus sco
18	459	28.7	273	1	PHAM_PHAVU	P15231 phaseolus v
19	452	28.3	233	1	LEC_VICFA	P02871 vicia faba
20	444.5	27.8	272	1	PHAL_PHAVU	P05087 phaseolus v
21	439	27.5	250	1	LEC1_LABAL	P23358 laburnum al
22	436.5	27.3	249	1	LEC2_ULEEU	P02973 ulex europe
23	435.5	27.2	285	1	LEC_SOYBN	P05046 glycine max
24	433	27.1	275	1	LEC1_DOLBI	P05045 dolichos bi
25	427.5	26.7	275	1	PHAE_PHAVU	P05088 phaseolus v
26	422	26.4	275	1	LEC5_DOLBI	P19588 dolichos bi
27	421.5	26.4	277	1	LEC1_WEDTR	Q01806 medicago tr
28	413	25.8	236	1	LEC_ONOVI	P02874 onobrychis
29	396	24.8	211	1	LEC_LENCU	P02870 lens culina
30	393.5	24.6	243	1	LEC4_GRISI	P24146 griffonia s
31	386	24.1	181	1	LECB_LATOC	P04122 lathyrus oc
32	383	24.0	290	1	LEC_BAUPU	P16030 baubinia pu
33	382.5	23.9	281	1	LEC_ERYCO	P16404 erythrina c

34 380 23.8 244 1 LEC_LATSP P16349 lathyrus sp
35 367.5 23.0 242 1 LEC1_PSOTE O24313 psophocarpu
36 365.5 22.9 237 1 CONA_CANLI P81460 canavalia l
37 364.5 22.8 236 1 LECA_CRAFL P81517 cratylia fl
38 363.5 22.7 237 1 CONA_CANBR P55915 canavalia b
39 363.5 22.7 237 1 CONA_CANVI P81461 canavalia v
40 360.5 22.5 237 1 LECA_DIOGR P08902 dioclea gra
41 359.5 22.5 237 1 LECA_DIOGU P08902 dioclea gra
42 359 22.5 273 1 LECG_ARAHY P01637 arachis gui
43 351.5 22.0 243 1 LEC1_ULEEU P02872 arachis hyp
44 350 21.9 265 1 LECN_PEA P22972 ulex europe
45 334 20.9 240 1 LEC_LOTTE P16270 pisum sativ
P19664 lotus tetra

ALIGNMENTS

RESULT 1

ID CONA_CANGL STANDARD; PRT: 290 AA.
AC P14894;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Concanavalin A precursor (Con A).
OS Canavalia gladiata (Sword bean) (Japanese jack bean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3824;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seedling;
RX MEDLINE=90127395; PubMed=2404793;
RA Yamauchi D., Minamikawa T.;
RT "Structure of the gene encoding concanavalin A from Canavalia
gladiata and its expression in Escherichia coli cells.";
RL FEBS Lett. 260:127-130(1990).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RA Yamauchi D., Nakamura K., Asahi T., Minamikawa T.;
RT "Nucleotide sequence of cDNA for concanavalin A from Canavalia
gladiata seeds";
RL Plant Cell Physiol. 30:147-150(1989).
CC -!- FUNCTION: D-MANNOSE SPECIFIC LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY
30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER
POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL
SEQUENCE. CLEAVAGE AFTER ASN AT POSITION 148, 163, AND 281 IS
FOLLOWED BY TRANSDISPOSITION AND LIGATION (BY FORMATION OF A NEW
PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X16041; CAA34163.1; --
DR PIR; A34139; A34139.
DR PIR; JQ2130; JQ2130.
DR HSP; P02866; I002.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.

DR pfam: PF00139; lectin_legB; 1.
 DR ProDom: PD000671; Lectin_legA; 1.
 DR ProDom: PD000711; Lectin_legB; 1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
 DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
 KW LECTIN; Calcium; Manganese; Glycoprotein; Signal.
 FT SIGNAL 1 29
 FT CHAIN 30 148 CONCAVALIN (SECOND PART).
 FT PROPEP 149 163 CONCAVALIN (FIRST PART).
 FT CHAIN 164 281
 FT PROPEP 282 290
 FT SITE 148 149 CLEAVAGE.
 FT SITE 163 164 CLEAVAGE.
 FT SITE 281 282 CLEAVAGE.
 FT METAL 171 171 MANGANESE (BY SIMILARITY).
 FT METAL 173 173 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 175 175 CALCIUM (BY SIMILARITY).
 FT METAL 177 177 CALCIUM (BY SIMILARITY).
 FT METAL 182 182 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 187 187 MANGANESE (BY SIMILARITY).
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (PROBABLE).
 SQ SEQUENCE 290 AA; 31421 MW; 3A1C9E9ADADA3580 CRC64;
 Query Match 41.2%; Score 658; DB 1; Length 290;
 Best Local Similarity 52.2%; Pred. No. 2.3e-44;
 Matches 140; Conservative 42; Mismatches 66; Indels 20; Gaps 9;
 QY 3 SLSFNTKFDLQKILFOGDATS-TNNVLQTLKDSGGNPGVSGRVLFSAPFHLWEN 61
 DB 36 ALHFHFNQFSKQDKLILQGDATGTGDNLELTRVSSNGSPQSSVGRALFYAPVHWES 95
 QY 62 SMVSSFTNLTQISTP--HPYAADGAFPLAHDVTIPPNWSGKFGILY--SNVERN 117
 DB 96 SAVVASFDATFTLKSPDSHP---ADGAFFTSNIDSSIPSGSTGRLLGLFPDANVIRN 152
 QY 118 SPTSE-NOSEFGDVNTDSRVVAVFEFTPNANI-DPNVRHIGIDVNSTKSKETARWQNG 175
 DB 153 STTIDFNAAAY---NADT-IVAVELDYPTDIDGPNYPHIGIDIKSVRSKTKAKWNG 208
 QY 176 KPTARISYNSAKSKSTVTFTYFGMEVVALSHDVLHAELPEWVRVGLSASTGEEKKNT 235
 DB 209 KVGTAHIYNSVKGRLSAVSVYPNGDSATVSYDVLNDNLVPEWVRVGLSASTGLYKETNT 268
 QY 236 ILSWFTSLKKNVKEPKEDMYIANV 263
 DB 269 ILSWFTSKLSNEIPD-----IATV 290
 RESULT 2
 ID CONA_CANEN STANDARD; PRT; 290 AA.
 AC P02866;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Concanavalin A precursor (Con A).
 OS Canavalia ensiformis (Jack bean) (Horse bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
 OX NCBI_TaxID=3823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=85086270; PubMed=3965973;
 RX Carrington D.M., Auffret A., Hanke D.E.;
 RT "Polypeptide ligation occurs during post-translational modification
 of concanavalin A.";
 RL Nature 313:64-67(1985).
 RN [2]
 RP SEQUENCE OF 30-148.
 RX MEDLINE=75095622; PubMed=1112813;
 RA Wang J.L., Cunningham B.A., Waxdal M.J., Edelman G.M.;
 RT "The covalent and three-dimensional structural of concanavalin A. I.
 Amino acid sequence of cyanogen bromide fragments F1 and F2.";
 J. Biol. Chem. 250:1490-1502(1975).
 RN [3]
 RP SEQUENCE OF 164-281.
 RX MEDLINE=75095623; PubMed=1112814;
 RA Cunningham B.A., Wang J.L., Waxdal M.J., Edelman G.M.;
 RT "The covalent and three-dimensional structure of concanavalin A. II.
 Amino acid sequence of cyanogen bromide fragment F3.";
 J. Biol. Chem. 250:1503-1512(1975).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
 RX MEDLINE=73053316; PubMed=4638345;
 RA Hardman K.D., Ainsworth C.F.;
 RT "Structure of concanavalin A at 2.4-A resolution.";
 Biochemistry 11:4910-4919(1972).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=75095624; PubMed=1112815;
 RA Becker J.W., Reeke G.N. Jr., Wang J.L., Cunningham B.A., Edelman G.M.;
 RT "The covalent and three-dimensional structure of concanavalin A. III.
 Structure of the monomer and its interactions with metals and
 saccharides.";
 J. Biol. Chem. 250:1525-1547(1975).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
 RA Parkin S., Rupp B., Hope H.;
 RT "Atomic resolution structure of concanavalin A at 120 K.";
 Acta Crystallogr. D 52:1161-1168(1996).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
 RX MEDLINE=20347885; PubMed=10748006;
 RA Bouckaert J., Dewalleer Y., Poortmans F., Wyns L., Loris R.;
 RT "The structural features of concanavalin A governing non-proline
 peptide isomerization.";
 J. Biol. Chem. 275:19778-19787(2000).
 CC -!- FUNCTION: D-mannose specific lectin.
 CC -!- SUBUNIT: Homotetramer.
 CC -!- PTM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY
 30-148. TO FORM A MATURE CHAIN THE PRECURSOR UNDERGOES FURTHER
 POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL
 SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS
 FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW
 PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
 AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
 SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOS LECTIN FAMILY.
 CC -!- DATABASE: NAME=Worthington enzyme manual;
 WWW="http://www.worthington-biochem.com/manual/C/CONA.html"
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X01632; CAA25787.1; -;
 DR PIR; A03357; CVJBP.
 DR PIR; A03358; CVJBP.
 DR PDB; 1CN1; 30-SEP-83.
 DR PDB; 2CNA; 31-JUL-94.
 DR PDB; 3CNA; 15-APR-91.
 DR PDB; 5CNA; 15-OCT-94.
 DR PDB; 1CON; 31-JAN-94.

DR PDB; 2CTV; 31-OCT-93.
 DR PDB; 1SCR; 15-OCT-94.
 DR PDB; 1SCS; 15-OCT-94.
 DR PDB; 1APN; 03-APR-96.
 DR PDB; 1CES; 15-FEB-97.
 DR PDB; 1CJP; 15-OCT-97.
 DR PDB; 1CVN; 14-OCT-96.
 DR PDB; 1ENQ; 17-AUG-96.
 DR PDB; 1ENR; 17-AUG-96.
 DR PDB; 1ENR; 16-FEB-99.
 DR PDB; 1ENS; 17-AUG-96.
 DR PDB; 1GIC; 20-AUG-97.
 DR PDB; 1JBC; 12-FEB-97.
 DR PDB; 1NLS; 26-NOV-97.
 DR PDB; 1ONL; 17-SEP-97.
 DR PDB; 1ONL; 17-SEP-97.
 DR PDB; 1TEI; 24-JUN-98.
 DR PDB; 1VAL; 11-JAN-97.
 DR PDB; 1VAM; 11-JAN-97.
 DR PDB; 1VLN; 01-APR-97.
 DR PDB; 1BXH; 07-OCT-98.
 DR PDB; 2CAU; 25-NOV-98.
 DR PDB; 2CAV; 25-NOV-98.
 DR PDB; 1DQ0; 19-JAN-00.
 DR PDB; 1DQ1; 19-JAN-00.
 DR PDB; 1DQ2; 19-JAN-00.
 DR PDB; 1DQ4; 19-JAN-00.
 DR PDB; 1DQ5; 19-JAN-00.
 DR PDB; 1DQ6; 19-JAN-00.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR LECTIN; Calcium; Manganese; Glycoprotein; Signal; 3D-structure.
 KW SIGNAL 1 29
 FT CHAIN 30 148 CONCAVALIN (SECOND PART).
 FT PROPEP 149 163
 FT CHAIN 164 281 CONCAVALIN (FIRST PART).
 FT PROPEP 282 290
 FT SITE 148 149 CLEAVAGE.
 FT SITE 163 164 CLEAVAGE.
 FT SITE 281 282 CLEAVAGE.
 FT METAL 171 171 MANGANESE.
 FT METAL 173 173 MANGANESE AND CALCIUM.
 FT METAL 175 175 CALCIUM.
 FT METAL 177 177
 FT METAL 182 182 MANGANESE.
 FT METAL 187 187
 FT METAL 187 187
 FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .).
 FT CONFLICT 33 33 E -> Q (IN REF. 2).
 FT CONFLICT 35 35 N -> D (IN REF. 2).
 FT CONFLICT 62 62 E -> D (IN REF. 3).
 FT CONFLICT 66 66 R -> E (IN REF. 3).
 FT CONFLICT 98 98 V -> T (IN REF. 3).
 FT CONFLICT 100 101 AS -> SA (IN REF. 3).
 FT CONFLICT 107 107 T -> A (IN REF. 3).
 FT STRAND 167 173
 FT STRAND 173 173
 FT HELIX 178 180
 FT TURN 181 181
 FT STRAND 187 192
 FT STRAND 199 202
 FT TURN 207 208
 FT STRAND 210 218
 FT TURN 219 222
 FT STRAND 223 229
 FT TURN 231 232
 FT STRAND 236 241
 FT HELIX 244 246
 FT TURN 247 247
 FT STRAND 251 259

FT STRAND 266 280
 FT TURN 283 284
 FT STRAND 286 289
 SQ SEQUENCE 290 AA; 31521 MW; 66CD1C62201720DD CRC64;
 Query Match 41.0%; Score 656; DB 1; Length 290;
 Best Local Similarity 52.2%; Pred. No. 3.2e-44;
 Matches 140; Conservative 41; Mismatches 67; Indels 20; Gaps 9;
 QY 3 SLSENETKFDLDKDLIFQGDATS-TNNVLQTLKLDGGNPGVGSYGRVLF\$APFHLWEN 61
 DB 36 ALHFENQFSKQKDLILQGDATTGTEGNRLTRVSSNGSPQGSVGRALFYAPVHIWES 95
 QY 62 SNAVSSFETNLTQISTP--HPYYAAGDAFAFLAPHDVTIPPNWKGKFLGLY--SNVFRN 117
 DB 96 SAVVASFEATFTFLIKSPDSHP---ADGIAFFISNDSSIPSGSTGRLLGLFPDANVIRN 152
 QY 118 SPTSE-NQSEFGDVNTDSRVVAVFEOTFPNANI-DPNVRHIGIDVNSTKSKETARWEMQNG 175
 DB 153 STTIDFNAAY---NADT-IVAVELDTYPTNDIGDPSYPHIGIDIKSVRSKKTAKNMONG 208
 QY 176 KTATARISYNSAKSKSTVTTFYPGMEVVALSHDVLHAELEPEWVRVGLSASTGEEKQKNT 235
 DB 209 KVGTAHIIVNSVDKRLSAVVSYPNADSATVSDVDLDNLVPEWVRVGLSASTGLYKETNT 268
 QY 236 IISWSFTSSLKKNVEKPEKEDMYIANVV 263
 DB 269 ILSWSFTSKLKSNEIPD-----IATVV 290
 RESULT 3
 ID LECA_DOLLA STANDARD; PRT; 237 AA.
 AC P38662;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE Lectin.
 OS Dolichos lab lab (Field bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
 RN NCBI_TaxID=35936;
 [1]
 RP SEQUENCE.
 RC STRAIN=cv. Lignosus; TISSUE=Seed;
 RA MEDLINE=94308133; PubMed=8034631;
 RA Gowda L.R., Savithri H.S., Rajagopal Rao D.;
 RT "The complete primary structure of a unique mannose/glucose-specific
 RL lectin from field bean (Dolichos lab lab).";
 J. Biol. Chem. 269:18789-18793(1994).
 CC -1- FUNCTION: D-MANNOSE/D-GLUCOSE-BINDING LECTIN. REQUIRES CA2+ AND
 CC M2+ IONS FOR FULL ACTIVITY.
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains.
 CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 DR HSSP; P02866; 10NA.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR LECTIN; Calcium; Manganese; Glycoprotein; Acetylation.
 FT CHAIN 1 105 BETA CHAIN.
 FT CHAIN 106 237 ALPHA CHAIN.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT MOD_RES 106 106 ACETYLATION.
 FT METAL 115 115 MANGANESE (BY SIMILARITY).
 FT METAL 117 117 MANGANESE AND CALCIUM (BY SIMILARITY).
 FT METAL 120 120 CALCIUM (BY SIMILARITY).
 FT METAL 122 122 CALCIUM (BY SIMILARITY).

```

FT METAL 127 127 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 132 132 MANGANESE (BY SIMILARITY).
SQ SEQUENCE 237 AA; 25718 MW; BEB7E84DC2895327 CRC64;

Query Match
Best Local Similarity 40.7%; Score 650; DB 1; Length 237;
Matches 143; Conservative 30; Mismatches 54; Indels 26; Gaps 7;

QY 1 AQSLSFNFTKFDLQKDLIFQGDATSTNNVLQTLKDSGGNPGVGSVGRVLFSAFFHLW 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 AQSLSFSFTKFDPNQDLIFQGTATS-----KLSAGNPVSSAGRVLYSAPLRWE 52
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NSMAVSSPFTNLTIQISTPHPYAADGFAFFLAPHDTVIPPNKSGKFLGLYSNVFRSPT 120
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 53 DSAVLTSPDP--TIYIFTNTSRKADGLA-FIAPPDSVI--SYHGGLGLFPNAESG-- 105
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 SENQSGDVTNDSRVVAVVEFDT-FPNANI-DPNYRHIGIDVNSIKSKETARWEWQNGKTA 178
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 -----IAESNVAVVEFDTYLPNDYGDPNYIHIGIDVNSIRSKVTASWDWQNGKIA 156
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 179 TARISYNSASKSTVTTFPGMEVVALSHVDVLAHELPEWVRVGLSASTGEEKKNTIIS 238
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 TAHISYNSYKSLSVTYYPYGRKGPATSYDIELHTVLPWVRVGLSASTGQNIERTVHS 216
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 239 WSFTSLKNEVK 251
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 217 WSFTSLWTNVAK 229
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 4
LEC_BOWMI STANDARD; PRT; 240 AA.
AC P42088;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Lectin (Agglutinin) (BMA).
OS Bowringia milderadii.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Bowringia.
OX NCBI_TaxID=28956;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=93385179; PubMed=8373823;
RA Chawla D., Animeshaun T., Hughes R.C., Harris A., Aitken A.;
RT "Bowringia milderadii agglutinin: polypeptide composition, primary
RT structure and homologues with other legume lectins.";
RL Biochim. Biophys. Acta 1202:38-46(1993).
CC -!- FUNCTION: BINDS PREFERENTIALLY TO OLIGOSACCHARIDES BEARING THE
CC SEQUENCE MAN-ALPHA-1->2 MAN-ALPHA-1->6 MAN-ALPHA-1->6MAN FOUND
CC IN EARLY STEPS OF GLYCOPROTEIN PROCESSING IN THE ENDOPLASMIC
CC RETICULUM. IT BINDS WEAKLY TO HIGHLY PROCESSED OLIGOSACCHARIDE
CC STRUCTURES.
CC -!- COFACTOR: CALCIUM AND MANGANESE ARE ESSENTIAL FOR THE SACCHARIDE-
CC BINDING AND CELL-AGGLUTININATING ACTIVITIES.
CC -!- SUBUNIT: HETEROTETRAMER OF TWO ALPHA AND TWO BETA CHAINS;
CC DISULFIDE BOND LINKED.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
HSSP; P02866; I0Q2.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; FALSE_NEG.
KW Lectin; Calcium; Manganese.
FT CHAIN 1 116 BETA CHAIN.
FT CHAIN 117 240 ALPHA CHAIN.
FT METAL 127 127 MANGANESE (BY SIMILARITY).

```

```

FT METAL 129 129 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 131 131 CALCIUM (BY SIMILARITY).
FT METAL 133 133 CALCIUM (BY SIMILARITY).
FT METAL 138 138 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 143 143 MANGANESE (BY SIMILARITY).
FT DISULFID 5 5 INTERCHAIN.
FT VARIANT 143 143 H -> S.
FT VARIANT 158 158 K -> R.
FT VARIANT 177 177 S -> R.
FT VARIANT 187 187 P -> V.
FT VARIANT 203 203 V -> A.
FT UNSURE 86 90
FT UNSURE 116 116
SQ SEQUENCE 240 AA; 25424 MW; 5A9F7FAF3A09B060 CRC64;

Query Match
Best Local Similarity 39.4%; Score 629.5; DB 1; Length 240;
Matches 133; Conservative 33; Mismatches 62; Indels 23; Gaps 5;

QY 1 AQSLSFNFTKFDLQKDLIFQGDAT-STNNVLQTLKDSGGNPGVGSVGRVLFSAFFHLW 59
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 ANSVCFPTDFESGQDLIFQGDASVGNKALQTLKVDKSGNPGQGSVGRALYTAPIRLW 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 ENSMAVSSPFTNLTIQI---STPHPYAADGFAFFLAPHDTVIPPNKSGKFLGLYSNVF 115
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 QSSSLVASFEFTFTFSISQGSSTP-----AAALTFIASPDTKIPSGSGRLGLFG---- 112
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 116 RNSPTSENQSGDVTNDSRVVAVVEFDTFPPNANI-DPNYRHIGIDVNSIKSKETARWEWQ 174
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 113 -----SSNAGSDNGVAVVEFDTYPTNDIGDPYRHIGIDVNSIRSKAASKWDWQ 163
   :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 175 GKTATARISYNSASKSTVTTFPGMEVVALSHVDVLAHELPEWVRVGLSASTGEEKQKN 234
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 GKTATAHISYNSASKRLSVSSYPNVVVSFDELNNVGPDPVVRVGSFATTTGYTTQTN 223
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 235 TIISWSFTSSL 245
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 224 NILAWSFRSSL 234
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
LEC2_CLALU STANDARD; PRT; 290 AA.
AC Q39529;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Agglutinin II precursor (CLAII) (LecCLAII).
OS Cladrastis lutea (Yellow wood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.
OX NCBI_TaxID=38412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-54.
RC TISSUE=Bark;
RX MEDLINE=96123235; PubMed=8534854;
RA van Damme E.J.M., Barre A., Bemer V., Rouge P., van Leuven F.,
RA Peumans W.J.;
RT "A lectin and a lectin-related protein are the two most prominent
RT proteins in the bark of yellow wood (Cladrastis lutea).";
RL Plant Mol. Biol. 29:579-598(1995).
CC -!- FUNCTION: MANNANOSE/GLUCOSE BINDING BARK LECTIN.
CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
CC SACCHARIDE-BINDING AND CELL-AGGLUTININATING ACTIVITIES (BY
CC SIMILARITY).

```


SQ SEQUENCE 270 AA; 29314 MW; 1FD655A2C4E550B3 CRC64;

Query Match 30.8%; Score 493; DB 1; Length 270;
Best Local Similarity 42.2%; Pred. No. 1.6e-31;
Matches 116; Conservative 53; Mismatches 80; Indels 26; Gaps 10;

OY 1 AQLSFFNFTFDDLDKDLIFQGDG-TSTNNVLQLTUKDGGGNPGVSGVRVLFPSAPHLW 59
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
16 AEILSFSPFVNSQEDLLQGDALVSSEGLQLTIVEN-GVPVNSTGRALYTAPVHIW 74
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
OY 60 ENSMA-VSSSETNLTIQISTPHPHYAADGFAFFLAPHDTVTIPPNWSCKFLGLYSNVFRNS 118
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
DB 75 DNSTGRVASATSFVFVKAPVASKSDAGTAFFLAPLNQI-HGAGGLYLGFNS----- 128
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
OY 119 PTSENQSGVDNTDSRVAVAEEDFPFNANIDPNRHHICIDVNSIKSKETARWEONGKTA 178
| : | : : : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
DB 129 -SSVSSSY-----QIVAVEEDHTNA-WDPNRHIGIDVNSVKSTKTWTGWNGEVA 179
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
OY 179 TARISYNASKKSTVTTFYFGMEV-VALSHDVLUHABELPEWRVGLSASTG---EEKOKN 234
| : | : : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
DB 180 NVLIYYQAATEMLTVSLTSPNOTSYTILSAVDLSKLTPENVRVGFTATTGLTTQYVEIN 239
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
OY 235 TITSWFTSSLKNEVKEPKREDMYIANVRVSWTWI 269
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :
DB 240 DVLSWSETSLTSDCGAEDNNVHLA-----SYAFI 270
| : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| : | : ||| :

RESULT 12

LCSL_BOBPS STANDARD; PRT; 285 AA.

ID LCSL_BOBPS STANDARD; PRT; 285 AA.

AC Q41162;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE Seed agglutinin I precursor (RPSAI) (LECRPAS1).

OS Robinia pseudoacacia (Black locust).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
ON NCBI_TaxID=35938;
RX [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-51.

RZ TISSUE=Seed;

RC MEDLINE=96191285; PubMed=8616218;

RA van Damme E.J.M., Barre A., Rouge P., van Leuven F., Peumans W.J.;
RT "The seed lectins of black locust (*Robinia pseudoacacia*) are encoded
RT by two genes which differ from the bark lectin genes.";
RL Plant Mol. Biol. 29:1197-1210(1995).
CC -!- FUNCTION: SEED LECTIN.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SEED.
CC -!- PTM: THREE OF THE FOUR POTENTIAL GLYCOSYLATION SITES ARE OCCUPIED.
CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licensee@isb-sib.ch).
CC -----
DR EMBL; U24250; AAC49272.1; --
DR HSSP; P19588; 1LUL,
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
KW

```

DR EMBL; U24249; AAC49271.1; -
DR HSSP; P19588; 1LUL.
DR InterPro; IPR000985; LECTIN_LEG4.
DR InterPro; IPR001220; LECTIN_LEG4.
DR Pfam; PF00138; lectin_leg4; 1.
DR Pfam; PF00139; lectin_leg4; 1.
DR ProDom; PD000671; LECTIN_LEG4; 1.
DR ProDom; PD000711; LECTIN_LEG4; 1.
DR PROSITE; PS00307; LECTIN_LEG4_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEG4_ALPHA; 1.
DR PROSITE; PS00308; LECTIN_LEG4_ALPHA; 1.
KW LECTIN; Calcium; Manganese; Glycoprotein; Signal; Multigene family.
FT SIGNAL 1 31
FT CHAIN 32 285 SEED AGGLUTININ II.
FT METAL 156 156 MANGANESE (BY SIMILARITY).
FT METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 162 162 CALCIUM (BY SIMILARITY).
FT METAL 166 166 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 171 171 MANGANESE (BY SIMILARITY).
FT CARBOHYD 147 147 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 285 AA; 31021 MW; 2C0B3249620294DE CRC64;

Query Match 30.2%; Score 483.5; DB 1; Length 285;
Best Local Similarity 44.0%; Pred. No. 9.2e-31;
Matches 114; Conservative 37; Mismatches 85; Indels 23; Gaps 8;

QY 3 SLGSFNTKFDLQKDLIFOGDATSTN-NVLQTLKLDGSGNPVGASGRVLFSAFPHLWEN 61
DQ 34 SLGSFPPKAPNOPLYIFORDALVTSTGVQLTNV-VNGVPSRKSLSGRALYAAPFQIWD 92
QY 62 SMA-VSSFTNLTIQISTPHYYAAGFAFFLAPHTVIPPNSWGKFLGLYSNVFRNSPT 120
DQ 93 TTGNVASFTSFIIQAPNPATTDGLAFPLAVDT--QPLDGLGMLGTFKN----- 143
QY 121 SENQSGDVNTDSRVAVAFEDFPNANIDPNYRHIGIDVNSIKSKETARWONGKTATA 180
DQ 144 -----GYENKSNQIVAFEDFENRHWDFGRHGINVNSIVSKIVPNWANGEVANV 197
QY 181 RTSYNSAKSKSTVTTFYPGMEVVALSHD-VDLHAELPEWVRVGLSASTGEEK---QKNTI 236
DQ 198 FTSYASTKSLASLYPSLETFSFIIHAIVDKRVLPWVRVGFSAATGDTGYVQTNV 257
QY 237 ISWSTSSLK-NNEVKEPK 254
DQ 258 LSWSFESNLPGGNSVASVK 276

RESULT 14
LECS_SOPJA STANDARD; PRT; 292 AA.
AC P93535;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Seed lectin precursor (LECS7ASG).
OS Sophora japonica (Japanese pagoda tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.
OX NCBI_TaxID=3897;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=97201486; PubMed=9049272;
RA van Damme E.J., Barre A., Rouge P., Peumans W.J.;
RT "Molecular cloning of the bark and seed lectins from the Japanese
RT pagoda tree (Sophora japonica).";
RL Plant Mol. Biol. 33:523-536(1997).
CC -!- FUNCTION: MANNOSYLGLUCOSE-SPECIFIC LECTIN.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U63011; AAB51441.1; -
DR HSSP; P19588; 1LUL.
DR GlycoSuiteDB; P93535; -
DR InterPro; IPR000985; LECTIN_LEG4.
DR InterPro; IPR001220; LECTIN_LEG4.
DR Pfam; PF00138; lectin_leg4; 1.
DR Pfam; PF00139; lectin_leg4; 1.
DR ProDom; PD000671; LECTIN_LEG4; 1.
DR ProDom; PD000711; LECTIN_LEG4; 1.
DR PROSITE; PS00307; LECTIN_LEG4_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEG4_ALPHA; 1.
DR PROSITE; PS00308; LECTIN_LEG4_ALPHA; 1.
KW LECTIN; Calcium; Manganese; Glycoprotein; Signal.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 292 SEED LECTIN.
FT METAL 163 163 MANGANESE (BY SIMILARITY).
FT METAL 165 165 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 167 167 CALCIUM (BY SIMILARITY).
FT METAL 169 169 CALCIUM (BY SIMILARITY).
FT METAL 172 172 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 177 177 MANGANESE (BY SIMILARITY).
FT CARBOHYD 82 82 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 292 AA; 31656 MW; A7431C29117A503E CRC64;

Query Match 30.0%; Score 480; DB 1; Length 292;
Best Local Similarity 40.7%; Pred. No. 1.8e-30;
Matches 112; Conservative 50; Mismatches 87; Indels 26; Gaps 9;

QY 1 AQLSFFNTKFDLQKDLIFOGDA-TSTNNVLQTLKLDGSGNPVGASGRVLFSAFPHLW 59
DQ 38 AEILSFPKPFASQDGLLQDGLVSSKGLQLTTVN-GVPIWNSTGRALYAPVHIW 96
QY 60 ENSMA-VSSFTNLTIQISTPHYYAAGFAFFLAPHTVIPPNSWGKFLGLYSNVFRNS 118
DQ 97 DKSTGRVASFATSFVVKAPVASKSADGIAFLAPPNNQI-QGPGGGHGLFHS----- 150
QY 119 PTSENQSGDVNTDSRVAVAFEDFPNANIDPNYRHIGIDVNSIKSKETARWONGKTA 178
DQ 151 -----SGYNSQYIIIVDFDTHINA-WDPNTRHIGIDVNSIINSTKTVTWQNGEYA 201
QY 179 TARISYNSAKSKSTVTTFYPGMEV-VALSVDVLDLHAELPEWVRVGLSASTG---EEKOKN 234
DQ 202 NVLISYQAATETLVSLTYPSSQTSYIILSAVDLSILPEWVRVGFATATGLTTQVVEH 261
QY 235 TTIWSTSSLKNNNEVKEPKEDMYTANVVSYSYTI 269
DQ 262 DVLSWSTFTTLETGCGAKDDNVHLV----SYAFI 292

RESULT 15
LECS_MEDTR STANDARD; PRT; 280 AA.
AC Q01807;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Truncated lectin 2 precursor.
GN LEC2.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
OX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Jemalong;
RX MEDLINE=92379255; PubMed=1511126;

```

```

RA  Bauchowitz M.A., Barker D.G., Nadaud I., Rouge P., Lescure B.;
RT  "Lectin genes from the legume Medicago truncatula.";
RL  Plant Mol. Biol. 19:1011-1017(1992).
CC  -1- MISCELLANEOUS: LEC2 IS PROBABLY NON FUNCTIONAL, SINCE A FRAMESHIFT
CC  MUTATION LEADS TO PREMATURE TRANSLATION TERMINATION AFTER ONLY 98
CC  AA. THE SEQUENCE BELOW IGNORES THIS FRAMESHIFT MUTATION.
CC  -1- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
CC  AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
CC  SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES.
CC  -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC  -----
CC  THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X60387; CAA42938.1; -.
DR  HSSP; P04122; lLOE.
DR  InterPro; IPR000985; Lectin_legA.
DR  InterPro; IPR001220; lectin_legB.
DR  Pfam; PF00138; lectin_legA; 1.
DR  Pfam; PF00139; lectin_legB; 1.
DR  ProDom; PD000671; Lectin_legA; 1.
DR  ProDom; PD000711; Lectin_legB; 1.
DR  PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR  PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW  Lectin; Manganese; Calcium; Signal; Multigene family; Glycoprotein.
FT  SIGNAL 1 26
FT  CHAIN 27 280
FT  METAL 148 148 MANGANESE (BY SIMILARITY).
FT  METAL 150 150 MANGANESE AND CALCIUM (BY SIMILARITY).
FT  METAL 152 152 CALCIUM (BY SIMILARITY).
FT  METAL 154 154 CALCIUM (BY SIMILARITY).
FT  METAL 158 158 MANGANESE AND CALCIUM (BY SIMILARITY).
FT  METAL 170 170 MANGANESE (BY SIMILARITY).
FT  CARBOHYD 163 163 N-LINKED (GLCNAC...) (POTENTIAL).
FT  CARBOHYD 272 272 N-LINKED (GLCNAC...) (POTENTIAL).
SQ  SEQUENCE 280 AA; 30473 MW; DB68690AD8015E81 CRC64;

Query Match 29.9%; Score 478; DB 1; Length 280;
Best Local Similarity 42.7%; Pred. No. 2.4e-30;
Matches 112; Conservative 37; Mismatches 79; Indels 34; Gaps 7;

QY 1 AQLSLFNFTRFDLDKDLIFOGDA-TSTNNVLQLTKLDSGGNPVCASVGRVLFSAFFHLW 59
DB 27 AETTSFSTIKFVPPQKNLIFOGDKASTGKLEJSK-----AVKNSIGRALYSAPIHIW 80
QY 60 ENSM-AVSSFTNLTQISTPHYPYAADGFAFFLAPHDTVIPPNWSGKFLGLY-SNVFRN 117
DB 81 DSKTGSVANFQTTFITATNTYNVADGLAFFIADPTKPSIHGGYLGVDSDSKTYKK 140
QY 118 SPTSENSFGDVNTSRVAVVEFTFPNANIDPN-----YRHGIDVNSIKSKETARW 170
DB 141 S-----IQTVAVEIDTFYNAQWDPNPGNISSTGRHIGIDVNSIKSISTVPW 186
QY 171 EWQNGKTATARIYSASKSKSTVTTFYFGMEVVALSHDVLHAELEPWVRVGLSASTGEE 230
DB 187 SLENNKANVAIGNGATNVLSDVDEYPLRIHYTLSHVVPKLDVVPEWVRIGFSSSTGAE 246
QY 231 KQKNTIISWFSFTSL-----KNN 248
DB 247 YSAHDILSWFSFDKINLGFENN 268

```

Search completed: February 26, 2003, 16:52:56
 Job time : 9.70037 secs


```

Db 209 KVTGTAHIYNSVCKRLSNAVSYNGDSATVSYDVLNDVLPWVRVGLSASTGLYKETNT 268
QY 236 IISWSTSSLKNNVEKPEKEDMYIANVV 263
Db 269 ILSWSFTSKLKSNEIPD-----IATVV 290

RESULT 2
CWJB
concanavalin A precursor - jack bean
C:Species: Canavalia ensiformis (jack bean)
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 28-May-1999
C:Accession: A03357; A60780; A60848
R:Carlington, D.M.; Auffret, A.; Hanke, D.E.
Nature 313, 64-67, 1985
A:Title: Polypeptide ligation occurs during post-translational modification of concanavalin A
A:Reference number: A03357; MUID:85086270; PMID:3965973
A:Accession: A03357
A:Molecule type: mRNA
A:Residues: 1-290 <CAR>
A:Cross-references: GB:X01632; NID:g312382; PIDN:CAA25787.1; PID:g17979
R:Chrispeels, M.J.; Hartl, P.M.; Sturm, A.; Faye, L.
J. Biol. Chem. 261, 10021-10024, 1986
A:Title: Characterization of the endoplasmic reticulum-associated precursor of concanavalin A
A:Reference number: A60780; MUID:86278043; PMID:3733700
A:Accession: A60780
A:Molecule type: protein
A:Residues: 30-41;160-173 <CHR>
R:Bowles, D.J.; Marcus, S.E.; Pappin, D.J.C.; Findlay, J.B.C.; Eliopoulos, E.; Maycox, B.
J. Cell Biol. 102, 1284-1297, 1986
A:Title: Posttranslational processing of concanavalin A precursors in jackbean cotyledon
A:Reference number: A60848; MUID:86168475; PMID:3958046
A:Accession: A60848
A:Molecule type: protein
A:Residues: 30-41;160-173 <BOW>
C:Comment: The mature chain (see PIR:CVJB) consists of residues 164-281 followed by 30-1
the signal sequence; cleavage after asparagines at positions 148, 163, and 281 is followed by
C:Superfamily: plant lectin
C:Keywords: glycoprotein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:149-163/Domain: glycopeptide #status predicted <GLP>
F:148-149/Cleavage site: Asn-Val (unidentified proteinase) #status experimental
F:152/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:163-164/Cleavage site: Asn-Ala (unidentified proteinase) #status experimental
F:281-282/Cleavage site: Asn-Glu (unidentified proteinase) #status experimental

Query Match 41.0%; Score 656; DB 1; Length 290;
Best Local Similarity 52.2%; Pred. No. 1.2e-43;
Matches 140; Conservative 41; Mismatches 67; Indels 20; Gaps 9;

QY 3 SLISNFTKFDLDKDLIFOGDATS--TNNVLQTLKLDGSGNPVGCASVGRVLFSAFPHLWEN 61
Db 36 ALHFENFQSKDKDLILQGDATTGEGNLRTRVSSNGSPQGSVGRALFAPVHVES 95
QY 62 SMVSNFTNLTIQISTP--HPYYAAGDGAFFLAPHDVTIPPNSWGKFLGLY--SNVERN 117
Db 96 SAVVASFEATFLLKSPDSHP---ADGIAFFISNIDSSIPSGSTGRLLGLFPDANVIRN 152
QY 118 SPTSE-NGSGDVNTDSRVAVVEFTFPNANI--DPNRYRHIGIDVNSIKSKETARWEQNG 175
Db 153 STTIDFNAAY---NADT-IVAVELDTPTNTDIDGFSYPHIGIDIKSVRSKTKAKNNQNG 208
QY 176 KTATARIISNSAKSKSTVTTFPGVGEVVALSHDVLHAELPEWVRVGLSASTGEERKQNT 235
Db 209 KVGTAHIYNSVCKRLSNAVSYNGDSATVSYDVLNDVLPWVRVGLSASTGLYKETNT 268
QY 236 IISWSTSSLKNNVEKPEKEDMYIANVV 263
Db 269 ILSWSFTSKLKSNEIPD-----IATVV 290

RESULT 3
mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea
C:Species: Cladrastis lutea
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S66357; S66301; S72502
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995
A:Title: A lectin and a lectin-related protein are the two most prominent proteins in
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66357
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-290 <VAN>
A:Cross-references: EMBL:U21959
A:Experimental source: bark
A:Accession: S66301
A:Molecule type: protein
A:Residues: 36-46,'S',48,'NEA',52-54 <VAF>
R:van Damme, E.J.M.
submitted to the EMBL Data Library, March 1995
A:Reference number: S72502
A:Accession: S72502
A:Molecule type: mRNA
A:Residues: 1-153,'L',155-230 <VAV>
A:Cross-references: EMBL:U21959; NID:g1141758; PIDN:AAC49137.1; PID:g1141759
A:Experimental source: bark
C:Superfamily: plant lectin
C:Keywords: glycoprotein; lectin
F:1-35/Domain: signal sequence #status predicted <SIG>
F:36-290/Product: mannose/glucose-binding lectin CLAI #status experimental <MAT>

Query Match 38.8%; Score 620; DB 2; Length 290;
Best Local Similarity 49.8%; Pred. No. 7.7e-41;
Matches 136; Conservative 39; Mismatches 70; Indels 28; Gaps 8;

QY 1 AQLSFTNFTKFDLDKDLIFOGDA--TSTNNVLQTLKLDGSGNPVGCASVGRVLFSAFPHL 58
Db 36 SDSLSFTFDNFRPQDRDILQGDAKISSGGSLQTLTKTDSKGPVRSVGRALYITPLHL 95
QY 59 WENSM-AVSSFETNLTIQISTP--HPYYAAGDGAFFLAPHDVTIPPNSWGKFLGLYS--N 113
Db 96 WDSSTNRLASFTTFTFLSSPTNNP---GDGIAFFIAPPETTIPGSSGGLGLFSPDN 152
QY 114 VFRNSPTSENGSGDVNTDSRVAVVEFTFPNANIDPNYRHIGIDVNSIKSKETARWEQ 173
Db 153 AVNNSL-----NQIVAVEFTFVNNMNDPSHRHIGIDVNTIKSSATVWQRE 199
QY 174 NGKTATARIISNSAKSKSTVTTEYPM---EVALSHDVLHAELPEWVRVGLSASTGEE 230
Db 200 NGLSATQAISYNSDTKKLSVVSSYPNTQANEDYTVSYDVLNDVLPWVRVGFSGSTGGY 259
QY 231 KQKNTIISWSTSSLKNNVEKPEKEDMYIANVV 263
Db 260 VQNHILSWTFNSNLQSSRAK--KEDIYIKRYV 290

RESULT 4
S66356
mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea
C:Species: Cladrastis lutea
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C:Accession: S66356; S66300
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995
A:Title: A lectin and a lectin-related protein are the two most prominent proteins in
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66356
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-293 <VAN>
A:Cross-references: EMBL:U21958; NID:g1141756; PIDN:AAC49136.1; PID:g1141757
A:Experimental source: bark
A:Accession: S66300

```


RESULT 10
S62691
agglutinin II precursor - black locust
N:Alternate names: lectin
C:Species: Robinia pseudoacacia (black locust)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S62691; S62686
R:van damme, E.J.M.; Barre, A.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 1197-1210, 1995
A:Title: The seed lectins of black locust (robinia pseudoacacia) are encoded by two g

[illegible]

RESULT 15

JQ1981
 lectin II - Scotch broom
 N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II
 C:Species: Cytisus scoparius (Scotch broom)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
 C:Accession: JQ1981
 R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.
 J. Biochem. 112, 366-375, 1992
 A:Title: The primary structure of the Cytisus scoparius seed lectin and a carbohydrate-
 A:Reference number: JQ1981; MOID:93054441; PMID:1429525
 A:Accession: JQ1981
 A:Molecule type: protein
 A:Residues: 1-249 <KON>
 A:Experimental source: seed
 C:Superfamily: plant lectin
 C:Keywords: lectin

Query Match	28.7%; Score 459.5; DB 2;	Length 249;
Best Local Similarity	43.2%; Pred. No. 1.9e-28;	
Matches 112;	Conservative 43;	Mismatches 75; Indels 29; Gaps 12;
QY	1 AQSLSFNFTHFDLQDKLIIF-QGDAATST-NNVLQLTKLDSCGNPVGASGVRYLFSAAPHL	58
Dd	:: ::::: :: : : : : :	:
Dd	1 SEELSFSTFKTDQKNLLLEQRDALITPTKGLQTTVEN-GKPAAYSLGRALYSTPIHI	59
QY	59 WENSMA-VSPETNLTIQIS-TPHPYAA-DGFAPFLAHDVTVPNPSWNGKFGLGSNVF	115
Dd	: : : : : : : : : : : : : : : :	:
Dd	60 WKDSTGDSEAFSTFFSVIDSNPMTAATDCLGAFFLAADT--QPDSAGLYLG----	113
QY	116 RNSPTSBNQSGDVNTSDRVVAVFDFFPNANIIDPNVR-HIGIDVNSIKSETARWEQN	174
Dd	:: ::: : : : : : : : : :	:
Dd	114 -----EKDSSY---NSSNQIIVAEEDYYINSANDPQTNPHGIDVNTLKSKVSWSGFRN	165
QY	175 GKATARISNASKKSTVTTFYPGMVEV-----VALSHDVDLHAELPEWWVGVSASTGE	229
Dd	: : : : : : : : : : : : : :	:
Dd	166 GNATVTLITYPPSKSLVASLVPSGGTSKDTSYIIANSVDLKATVPPEWRIGFSATTQG	225
QY	230 EK----QKNNTIISFTSSL	245
Dd	: : :	:
Dd	226 TDNYIETHDIILWSFKSKL	244

Search completed: February 26, 2003, 16:51:32
Job time : 23.6966 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 13.5131 Seconds
(without alignments)
574.823 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFDPNQEDLIF.....LWTVAKKENKVIIRCVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	99.4	270	4	US-08-881-189B-2
2	1362	99.4	286	4	US-08-881-189B-23
3	575.5	42.0	132	4	US-08-881-189B-12
4	463	33.8	285	3	US-09-141-821-2
5	441	32.2	285	3	US-09-141-821-4
6	440	32.1	285	3	US-09-141-821-1
7	439	32.0	285	3	US-09-141-821-5
8	430.5	31.4	105	4	US-08-881-189B-13
9	420	30.7	262	2	US-08-038-761A-1
10	321.5	23.5	274	3	US-09-141-821-3
11	244	17.8	632	4	US-09-228-986-77
12	107.5	7.8	1095	4	US-09-206-942-43
13	107.5	7.8	1101	4	US-09-206-942-43
14	105	7.7	18	4	US-08-881-189B-14
15	101	7.4	304	1	US-07-851-976B-8
16	101	7.4	304	1	US-08-291-609-8
17	101	7.4	304	1	US-08-401-136-8
18	101	7.4	304	1	US-08-850-554-8
19	97	7.1	1876	2	US-08-609-049A-12
20	97	7.1	1876	2	US-08-609-049A-28
21	97	7.1	1876	2	US-09-170-996-12
22	97	7.1	1876	4	US-09-170-996-28
23	91	6.6	675	4	US-08-426-509A-4
24	91	6.6	675	5	PCT-US95-05008-4
25	91	6.6	775	1	US-07-603-133B-15
26	90	6.6	775	1	US-07-603-133B-16
27	89	6.5	724	3	US-09-121-964-1

28 89 6.5 860 1 US-08-092-817-4 Sequence 4, Appli
29 89 6.5 860 4 US-08-485-128-4 Sequence 4, Appli
30 89 6.5 860 4 US-09-804-778A-8 Sequence 8, Appli
31 89 6.5 1222 4 US-09-206-942-37 Sequence 37, Appli
32 89 6.5 1228 4 US-09-206-942-34 Sequence 34, Appli
33 89 6.5 1410 2 US-08-470-058-4 Sequence 4, Appli
34 89 6.5 1410 3 US-09-037-188-4 Sequence 4, Appli
35 89 6.5 1410 4 US-09-285-310-4 Sequence 8, Appli
36 86 6.3 790 3 US-08-537-361E-8 Sequence 4, Appli
37 85.5 6.2 463 2 US-08-853-659A-52 Sequence 52, Appli
38 84.5 6.2 616 4 US-09-298-367B-11 Sequence 11, Appli
39 84.5 6.2 1250 1 US-08-441-139-9 Sequence 9, Appli
40 84 6.1 790 4 US-08-817-707-8 Sequence 8, Appli
41 84 6.1 1002 4 US-09-268-347-24 Sequence 24, Appli
42 83 6.1 1104 4 US-09-268-347-28 Sequence 28, Appli
43 83 6.1 1104 4 US-09-268-347-34 Sequence 34, Appli
44 82.5 6.0 514 3 US-08-467-023-134 Sequence 134, App
45 82.5 6.0 853 4 US-08-913-880C-17 Sequence 17, Appli

ALIGNMENTS

RESULT 1
US-08-881-189B-2
; Sequence 2 Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-2

Query Match 99.4%; Score 1362; DB 4; Length 270;
Best Local Similarity 99.6%; Pred. No. 4.9e-128;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNNVLQTKLDSAGNPVSSSAGRVLYSAPRLWE 60

Db 1 AQSLSFSTKFDPNQEDLIFQGHATSTNNVLQTKLDSAGNPVSSSAGRVLYSAPRLWE 60

61	Qy	DSAVLTSFTDIINFEISTPYTSRIADGLAFFIAPDPSVSYHGGFLGLFPNNANTLNNSST	120
61	Db	DSAVLTSFTDIINFEISTPYTSRIADGLAFFIAPDPSVSYHGGFLGLFPNNANTLNNSST	120
121	Qy	SENQTTTKAASSNVAVFDTYLPNDYGDPNYIHIGIDVNSIRSKVATAKDWQNGKIATA	180
121	Db	SENQTTTKAASSNVAVFDTYLPNDYGDPNYIHIGIDVNSIRSKVATAKDWQNGKIATA	180
181	Qy	HISYNSVKRLSVTSYIAGSKPATLSYDIELHTVLPFWVRVGLSASTGDKERNIVHWS	240
181	Db	HISYNSVKRLSVTSYIAGSKPATLSYDIELHTVLPFWVRVGLSASTGDKERNIVHWS	240
241	Qy	FTSSLWNTNVAKKENENKYITRGVL	264
241	Db	FTSSLWNTNVAKKENENKYITRGVL	264

RESULT 2
US-08-881-189B-23
Sequence 23, Application US/08881189B
Patent No. 6310195
GENERAL INFORMATION:
APPLICANT: Colucci et al.
TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881,189B
FILING DATE: June 24, 1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: 381,44 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-881-189B-23

Db	143	SENQTTKARSSNVAFEDTYLNPIDGPNYIHIGIDVNSIRSKVTAKWDQNGKIATA	202
Qy	181	HISYNSVKRLSTSYTYAGSKPATLSYDIELHTVLPFWVRVGLSASTGQDKERNTVHSWS	240
Db	203	HISYNSVKRLSTSYTYAGSKPATLSYDIELHTVLPFWVRVGLSASTGQDKERNTVHSWS	262
Qy	241	FTSSLWTNVAKKENNKYITRGVL	264
Db	263	FTSSLWTNVAKKENNKYITRGVL	286

```

1  RESULT 3
2  US-08-881-189B-12
3  ; Sequence 12, Application US/08881189B
4  ; Patent No. 6310195
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Colucci et al.
7  ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
8  ; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
9  ; NUMBER OF SEQUENCES: 24
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSEE: Hoffmann & Baron, LLP
12 ; STREET: 350 Jericho Turnpike
13 ; CITY: Jericho
14 ; STATE: New York
15 ; COUNTRY: USA
16 ; ZIP: 11753
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
19 ; COMPUTER: IBM compatible
20 ; OPERATING SYSTEM: MS-DOS
21 ; SOFTWARE: Wordperfect
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/881,189B
24 ; FILING DATE: June 24, 1997
25 ; CLASSIFICATION: 424
26 ; PRIOR APPLICATION DATA:
27 ; APPLICATION NUMBER:
28 ; FILING DATE:
29 ; ATTORNEY/AGENT INFORMATION:
30 ; NAME: Feit, Irving N.
31 ; REGISTRATION NUMBER: 28,601
32 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
33 ; TELECOMMUNICATION INFORMATION:
34 ; TELEPHONE: (516) 822-3550
35 ; TELEFAX: (516) 822-3582
36 ; INFORMATION FOR SEQ ID NO: 12:
37 ; SEQUENCE CHARACTERISTICS:
38 ; LENGTH: 132 amino acids
39 ; TYPE: amino acid
40 ; TOPOLOGY: linear
41 ; MOLECULE TYPE: peptide
42 ;
43 ;
44 ;
45 ;
46 ;
47 ;
48 ;
49 ;
50 ;
51 ;
52 ;
53 ;
54 ;
55 ;
56 ;
57 ;
58 ;
59 ;
60 ;
61 ;
62 ;
63 ;
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;
101 ;
102 ;
103 ;
104 ;
105 ;
106 ;
107 ;
108 ;
109 ;
110 ;
111 ;
112 ;
113 ;
114 ;
115 ;
116 ;
117 ;
118 ;
119 ;
120 ;
121 ;
122 ;
123 ;
124 ;
125 ;
126 ;
127 ;
128 ;
129 ;
130 ;
131 ;
132 ;
133 ;
134 ;
135 ;
136 ;
137 ;
138 ;
139 ;
140 ;
141 ;
142 ;
143 ;
144 ;
145 ;
146 ;
147 ;
148 ;
149 ;
150 ;
151 ;
152 ;
153 ;
154 ;
155 ;
156 ;
157 ;
158 ;
159 ;
160 ;
161 ;
162 ;
163 ;
164 ;
165 ;
166 ;
167 ;
168 ;
169 ;
170 ;
171 ;
172 ;
173 ;
174 ;
175 ;
176 ;
177 ;
178 ;
179 ;
180 ;
181 ;
182 ;
183 ;
184 ;
185 ;
186 ;
187 ;
188 ;
189 ;
190 ;
191 ;
192 ;
193 ;
194 ;
195 ;
196 ;
197 ;
198 ;
199 ;
200 ;
201 ;
202 ;
203 ;
204 ;
205 ;
206 ;
207 ;
208 ;
209 ;
210 ;
211 ;
212 ;
213 ;
214 ;
215 ;
216 ;
217 ;
218 ;
219 ;
220 ;
221 ;
222 ;
223 ;
224 ;
225 ;
226 ;
227 ;
228 ;
229 ;
230 ;
231 ;
232 ;
233 ;
234 ;
235 ;
236 ;
237 ;
238 ;
239 ;
240 ;
241 ;
242 ;
243 ;
244 ;
245 ;
246 ;
247 ;
248 ;
249 ;
250 ;
251 ;
252 ;
253 ;
254 ;
255 ;
256 ;
257 ;
258 ;
259 ;
260 ;
261 ;
262 ;
263 ;
264 ;
265 ;
266 ;
267 ;
268 ;
269 ;
270 ;
271 ;
272 ;
273 ;
274 ;
275 ;
276 ;
277 ;
278 ;
279 ;
280 ;
281 ;
282 ;
283 ;
284 ;
285 ;
286 ;
287 ;
288 ;
289 ;
290 ;
291 ;
292 ;
293 ;
294 ;
295 ;
296 ;
297 ;
298 ;
299 ;
300 ;
301 ;
302 ;
303 ;
304 ;
305 ;
306 ;
307 ;
308 ;
309 ;
310 ;
311 ;
312 ;
313 ;
314 ;
315 ;
316 ;
317 ;
318 ;
319 ;
320 ;
321 ;
322 ;
323 ;
324 ;
325 ;
326 ;
327 ;
328 ;
329 ;
330 ;
331 ;
332 ;
333 ;
334 ;
335 ;
336 ;
337 ;
338 ;
339 ;
340 ;
341 ;
342 ;
343 ;
344 ;
345 ;
346 ;
347 ;
348 ;
349 ;
350 ;
351 ;
352 ;
353 ;
354 ;
355 ;
356 ;
357 ;
358 ;
359 ;
360 ;
361 ;
362 ;
363 ;
364 ;
365 ;
366 ;
367 ;
368 ;
369 ;
370 ;
371 ;
372 ;
373 ;
374 ;
375 ;
376 ;
377 ;
378 ;
379 ;
380 ;
381 ;
382 ;
383 ;
384 ;
385 ;
386 ;
387 ;
388 ;
389 ;
390 ;
391 ;
392 ;
393 ;
394 ;
395 ;
396 ;
397 ;
398 ;
399 ;
400 ;
401 ;
402 ;
403 ;
404 ;
405 ;
406 ;
407 ;
408 ;
409 ;
410 ;
411 ;
412 ;
413 ;
414 ;
415 ;
416 ;
417 ;
418 ;
419 ;
420 ;
421 ;
422 ;
423 ;
424 ;
425 ;
426 ;
427 ;
428 ;
429 ;
430 ;
431 ;
432 ;
433 ;
434 ;
435 ;
436 ;
437 ;
438 ;
439 ;
440 ;
441 ;
442 ;
443 ;
444 ;
445 ;
446 ;
447 ;
448 ;
449 ;
450 ;
451 ;
452 ;
453 ;
454 ;
455 ;
456 ;
457 ;
458 ;
459 ;
460 ;
461 ;
462 ;
463 ;
464 ;
465 ;
466 ;
467 ;
468 ;
469 ;
470 ;
471 ;
472 ;
473 ;
474 ;
475 ;
476 ;
477 ;
478 ;
479 ;
480 ;
481 ;
482 ;
483 ;
484 ;
485 ;
486 ;
487 ;
488 ;
489 ;
490 ;
491 ;
492 ;
493 ;
494 ;
495 ;
496 ;
497 ;
498 ;
499 ;
500 ;
501 ;
502 ;
503 ;
504 ;
505 ;
506 ;
507 ;
508 ;
509 ;
510 ;
511 ;
512 ;
513 ;
514 ;
515 ;
516 ;
517 ;
518 ;
519 ;
520 ;
521 ;
522 ;
523 ;
524 ;
525 ;
526 ;
527 ;
528 ;
529 ;
530 ;
531 ;
532 ;
533 ;
534 ;
535 ;
536 ;
537 ;
538 ;
539 ;
540 ;
541 ;
542 ;
543 ;
544 ;
545 ;
546 ;
547 ;
548 ;
549 ;
550 ;
551 ;
552 ;
553 ;
554 ;
555 ;
556 ;
557 ;
558 ;
559 ;
560 ;
561 ;
562 ;
563 ;
564 ;
565 ;
566 ;
567 ;
568 ;
569 ;
570 ;
571 ;
572 ;
573 ;
574 ;
575 ;
576 ;
577 ;
578 ;
579 ;
580 ;
581 ;
582 ;
583 ;
584 ;
585 ;
586 ;
587 ;
588 ;
589 ;
590 ;
591 ;
592 ;
593 ;
594 ;
595 ;
596 ;
597 ;
598 ;
599 ;
600 ;
601 ;
602 ;
603 ;
604 ;
605 ;
606 ;
607 ;
608 ;
609 ;
610 ;
611 ;
612 ;
613 ;
614 ;
615 ;
616 ;
617 ;
618 ;
619 ;
620 ;
621 ;
622 ;
623 ;
624 ;
625 ;
626 ;
627 ;
628 ;
629 ;
630 ;
631 ;
632 ;
633 ;
634 ;
635 ;
636 ;
637 ;
638 ;
639 ;
640 ;
641 ;
642 ;
643 ;
644 ;
645 ;
646 ;
647 ;
648 ;
649 ;
650 ;
651 ;
652 ;
653 ;
654 ;
655 ;
656 ;
657 ;
658 ;
659 ;
660 ;
661 ;
662 ;
663 ;
664 ;
665 ;
666 ;
667 ;
668 ;
669 ;
670 ;
671 ;
672 ;
673 ;
674 ;
675 ;
676 ;
677 ;
678 ;
679 ;
680 ;
681 ;
682 ;
683 ;
684 ;
685 ;
686 ;
687 ;
688 ;
689 ;
690 ;
691 ;
692 ;
693 ;
694 ;
695 ;
696 ;
697 ;
698 ;
699 ;
700 ;
701 ;
702 ;
703 ;
704 ;
705 ;
706 ;
707 ;
708 ;
709 ;
710 ;
711 ;
712 ;
713 ;
714 ;
715 ;
716 ;
717 ;
718 ;
719 ;
720 ;
721 ;
722 ;
723 ;
724 ;
725 ;
726 ;
727 ;
728 ;
729 ;
730 ;
731 ;
732 ;
733 ;
734 ;
735 ;
736 ;
737 ;
738 ;
739 ;
740 ;
741 ;
742 ;
743 ;
744 ;
745 ;
746 ;
747 ;
748 ;
749 ;
750 ;
751 ;
752 ;
753 ;
754 ;
755 ;
756 ;
757 ;
758 ;
759 ;
760 ;
761 ;
762 ;
763 ;
764 ;
765 ;
76
```

RESULT 4

APPLICANT: Szuzanna Magdolna BARDOZ
APPLICANT: Richard Michael John PALMER
APPLICANT: Neil William RISH
APPLICANT: Gyorgy J. KOTELIS
TITLE OF INVENTION: LECTIN COMPOSITIONS AND USES THEREOF
FILE REFERENCE: 48656
CURRENT APPLICATION NUMBER: US/09/141,821
CURRENT FILING DATE: 1998-08-28
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 5
LENGTH: 285
TYPE: PRT
ORGANISM: Human
US-09-141-821-5

Query Match 32.0%; Score 439; DB 3; Length 285;
Best Local Similarity 42.5%; Pred. No. 7.le-36;
Matches 111; Conservative 39; Mismatches 87; Indels 24; Gaps 9;
QY 2 GSLSFSTKFDNPQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 60
DB 33 GSLSFSTKFDNPQEDLIFQGHATSTN-NVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 60
QY 61 D-SAVLTSTFTDIINFEISTPTTSRIADGLAFTIAPPDVSIVSYHGGFLGLFPNA 112
DB 92 STTGNVASFVTSFIIQAPNPATTADGLAFTIAPPDVSIVSYHGGFLGLFPNA 101
QY 120 TSENQTTKKAASSNVVAFEDYLPDYGDPNY IHGIDVNSTRSKVTAKWQNCKIAT 179
DB 149 -----SNQIVAFEDYFNSRHW--DPTRGRMGINVSIVSKTPVWNWANGEVAN 196
QY 180 AHISYNSVKRLSVTSYXAGSKPATLSYDI-ELHTVLPPEWVRVGLSASTGQDK---ERNT 235
DB 197 VFISYEASTKSLTASLVPSLETSEIHAIVDKVLPPEWVRVGLSASTGQDK---ERNT 235
QY 236 VHSWSTSSL--WTNVAKEN 254
DB 257 VLSWSFESNLPGGNSVASVKN 277

RESULT 8
US-08-881-189B-13
Sequence 13, Application US/08881189B
Patent No. 6310195
GENERAL INFORMATION:
APPLICANT: Colucci et al.
TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
PROGENITOR CELL PRESERVATION FACTOR
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881,189B
FILING DATE: June 24, 1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: 381-44 PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-881-189B-13

Query Match 31.4%; Score 430.5; DB 4; Length 105;
Best Local Similarity 83.9%; Pred. No. 1.le-35;
Matches 94; Conservative 0; Mismatches 7; Indels 11; Gaps 3;

QY 1 AGSLSFSTKFDNPQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 60
DB 1 AQSLSFSTKFDNPQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPRLWE 52
QY 61 DSAVLTSFTDIINFEISTPTTSRIADGLAFTIAPPDVSIVSYHGGFLGLFPNA 112
DB 53 DSAVLTSFTDIINFEISTPTTSRIADGLAFTIAPPDVSIVSYHGGFLGLFPNA 101

RESULT 9
US-08-038-761A-1
Sequence 1, Application US/08038761A
Patent No. 5945589
GENERAL INFORMATION:
APPLICANT: Rao, A. Gururaj; Kumar, M. Arun
TITLE OF INVENTION: Derivatives of Bauhinia
TITLE OF INVENTION: Purpurea Lectin and Their Use as Larvicides
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pioneer Hi-Bred International, Inc.
STREET: 700 Capital Square, 400 Locust
STREET: Street
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft WORKS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,761A
FILING DATE: 24-JUL-1992
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Roth, Michael J.
REGISTRATION NUMBER: 29,342
REFERENCE/DOCKET NUMBER: 0210 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 245-3594
TELEFAX: (515) 245-3634
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION:
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:


```
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-43

Query Match      7.8%; Score 107.5; DB 4; Length 1095;
Best Local Similarity 24.1%; Pred. No. 0.066;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

RESULT 12
US-09-206-942-45
; Sequence 45, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-45

Query Match      7.8%; Score 107.5; DB 4; Length 1095;
Best Local Similarity 24.1%; Pred. No. 0.066;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

; 26 STNNVQVTKLDSAGNPVSSSAGRVLYSAPLR---LWEDSAVLTSFDTIINFESTPYTS 82
; 230 STNSYLNT-IDNSGRSPGAGPLRRSLGICISFNNDTVFNVASGSAVNFISIKPPIVS 288
; 83 RTADGLAFIAPDPSVISYHGGFLGLFPNANTLN-----NSSSTENOT---TTK 128
; 289 NVHDGNHTLFGNVSVLG--GGDVNFHFENASSNHWTHGVVKSQNFNASEGSLRFKSE 346
; 129 AASNVAVEFDYTLNPDYGDPNYIHI-GIDVNSIRSKVTAK-WDWQNGKIATAHISYNS 186
; 347 GSTRTAFIESDLTLNATGNSLNQVAGIDGNLQKSLVANKNITFEGGNITLA----- 400
; 187 VSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGL-SASTGQDKERNTVHWSFTSSL 245
; 401 -----ADKKP-----IEIKGNITVKEGANVTLRSANYGNDKSALSIRG----- 438
; 246 WTNVAKKEN 254
; 439 --NVTNKN 445

RESULT 13
US-09-206-942-43
; Sequence 43, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; TITLE OF INVENTION: Molecular Weight Proteins
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 1101
```

```
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-206-942-43

Query Match      7.8%; Score 107.5; DB 4; Length 1101;
Best Local Similarity 24.1%; Pred. No. 0.067;
Matches 60; Conservative 37; Mismatches 99; Indels 53; Gaps 11;

; 26 STNNVQVTKLDSAGNPVSSSAGRVLYSAPLR---LWEDSAVLTSFDTIINFESTPYTS 82
; 236 STNSYLNT-IDNSGRSPGAGPLRRSLGICISFNNDTVFNVASGSAVNFISIKPPIVS 294
; 83 RIADGLAFIAPDPSVISYHGGFLGLFPNANTLN-----NSSSTENOT---TTK 128
; 295 NVHDGNHTLFGNVSVLG--GGDVNFHFENASSNHWTHGVVKSQNFNASEGSLRFKSE 352
; 129 AASNVAVEFDYTLNPDYGDPNYIHI-GIDVNSIRSKVTAK-WDWQNGKIATAHISYNS 186
; 353 GSTRTAFIESDLTLNATGNSLNQVAGIDGNLQKSLVANKNITFEGGNITLA----- 406
; 187 VSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGL-SASTGQDKERNTVHWSFTSSL 245
; 407 -----ADKKP-----IEIKGNITVKEGANVTLRSANYGNDKSALSIRG----- 444
; 246 WTNVAKKEN 254
; 445 --NVTNKN 451

RESULT 14
US-08-881-189B-14
; Sequence 14, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-14

Query Match      7.7%; Score 105; DB 4; Length 18;
```



```

RESULT 4
US-09-782-906-3
; Sequence 3, Application US/09782906
; Patent No. US20010051369A1
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Rittenhouse Pruss, Jennifer L.
; APPLICANT: Murphy, Dennis J.
; APPLICANT: Maffia III, Anthony M.
; APPLICANT: Bylina, Edward J.
; APPLICANT: Coleman, William J.
; TITLE OF INVENTION: Variant Galactose Oxidase, Nucleic Acid Encoding Same, And Met
; FILE OF INVENTION: Using Same
; FILE REFERENCE: HER-0040
; CURRENT APPLICATION NUMBER: US/09/782.906
; CURRENT FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/185,001
; PRIOR FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: No. US20010051369A1el Sequence
US-09-782-906-3

Query Match          7.3%; Score 99.5; DB 10; Length 639;
Best Local Similarity 23.7%; Pred. No. 0.33;
Matches 57; Conservative 43; Mismatches 90; Indels 51; Gaps 13;

QY 13 PNQEDILFQGHATSTNNVLQTKLDSAGNPVSSAGRGVLYSAPLRLWEDSAVLTSFDTH 72
| : : : | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
DB 403 PDYQD-----SDATNNAHIIITGEGTSPNTVFASNG--LHFA--RTFHTSVVLPGSGTFI 454

```


FILE REFERENCE: 0342/IG548-US2
CURRENT APPLICATION NUMBER: US/09/893,519A
PRIOR FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/215,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 41
LENGTH: 415
TYPE: PRT
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Corresponds to SEQ ID NO: 114
US-09-893-519A-41

Query Match 6.7%; Score 92; DB 9; Length 415;
Best Local Similarity 23.7%; Pred. No. 0.92;
Matches 54; Conservative 30; Mismatches 100; Indels 44; Gaps 9;

Qy 3 SISLFTKFPDQEDLIFQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPRLWEDS 62
Db 17 SISGQLKFDSDIAPYIKELTEKEN---VRKVPFSGNTIGIEASKAL-----SE 64
Qy 63 AVLTSEFDIINEFTPTYSRTIADGLAFPIAPDSVISYHGGFLGFLPNANTLNNSSTSE 122
Db 65 ALLKHKDTIVEINFSDLYTGRLENTET-----PQS-LEYLLPALSKLPNKLINLSDNAF 117
Qy 123 NOTT---TKAASNNVAVFDTYLPNDYDGNVYIHGIDV-----NSIRSKVT 167
Db 118 GLQTDPIEAYLAKAVSIEHLILNNGMPFAGSRIGGSLFKLAKAKAGKESLTKFIC 177
Qy 168 AKWDQWNGKIATAHISYNSVS-----KRLSVTSYYA-GSKPATLSYDIE 210
Db 178 GRNRLNGS-----VNYLSVGLRNHKKDLVRLYQNGIRPAGISKLIVE 220

RESULT 8
US-09-841-132-394
Sequence 394, Application US/09841132
Patent No. US20020061848A1
GENERAL INFORMATION:
APPLICANT: Bhatia, Ajay
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Probst, Peter
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
DIAGNOSIS OF CHLAMYDIAL INFECTION
FILE REFERENCE: 210121.469C8
CURRENT APPLICATION NUMBER: US/09/841,132
PRIOR FILING DATE: 2001-04-23
NUMBER OF SEQ ID NOS: 599
SOFTWARE: FastSeq for Windows Version 3.0/4.0
SEQ ID NO 394
LENGTH: 1723
TYPE: PRT
ORGANISM: Chlamydia pneumoniae
US-09-841-132-394

Query Match 6.7%; Score 91.5; DB 10; Length 1723;
Best Local Similarity 23.5%; Pred. No. 6.9;
Matches 63; Conservative 32; Mismatches 106; Indels 67; Gaps 12;

Qy 14 NQEDLIFQGHATSTNNVQVTKLDSAGNPVSSSA---GRVLYSAPRLWEDSAVLTSFDT 70
Db 1007 NPEITFKGNYATGGAIGCIDITNGSPPRKVSADNGSVLF-----QDNSALNRGA 1059
Qy 71 IINFEISTPTYSRTIADGLAFPIAPDSVISYHGG-----FLGLFPNANTL-NNSSTSE 122
Db 1060 IYGETIDISRTG-----ATFIGNS-----SKHDGSAICCSALTALPAPNSQLIFENNKVTE 1109
Qy 123 NQTTKAASNNVAVFDTYLPNDYDGNVYIHGIDVNSIRSKVTAKWDQWNGKI----- 177

Db 1110 TTATTKASINNLGAA-----IYG-----NNETSDVTISLSAENGSIFFKNN 1150
Qy 178 -ATAHISYNSVSKRLSVTSYAGSKPATLSVDIELHIVLPWVRVGLSASTGQDKERTV 236
Db 1151 LCTATNKYCSAGNVKFTAJEASAGKALSFYDA-----VNVSTKETNAQELKNE- 1200
Qy 237 HSNSTSSSLWTNVAKKENK-YITRGV 263
Db 1201 ---KATSTGTLFSGELHKNKSIPOKV 1225

RESULT 9
US-09-801-368-226
Sequence 226, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 226
LENGTH: 433
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-226

Query Match 6.6%; Score 91; DB 10; Length 433;
Best Local Similarity 23.0%; Pred. No. 1.2;
Matches 70; Conservative 37; Mismatches 108; Indels 90; Gaps 15;

Qy 24 ATSTNNVQVTKLDSAG-----NPVSSSAGRVLYSAPRLWEDSAVLTSFD 69
Db 8 ATSKNQVLQRRPLETSNHNHGFASSLQAIPEMTMSGSDNASFQSLPLSM--SSSQSTSS 65
Qy 70 TIINFEISTP-YTSRIADGL-AFFIAPDPSVISYHGGFLG-----LFPNAN 113
Db 66 RRENPNAPPEYTDRADEIKRLLASSPSRRSHSSSMHSASRRSSVAESGSLSDNAS 125
Qy 114 TLNNS-----STSENQTTTKAASN-----VVAVEFDYTL-----NPD--YGDPNY 152
Db 126 SYOSSIFSAPSTVHTQLTNDSSFEFPNKLITRVSLEALPKTFYDMYSPDILLADPSN 185
Qy 153 I-----HIGIDVNSIRS-----KVAKWDQWNGKIATAHISYNSVSKRLSVTS 195
Db 186 ILCNGRPKFTKRELDWDLNDLRSLLIVLEKLRPEWGNQLPEVIT--VGNMPPQRLQLLP 243
Qy 196 YVAGSKPATLSYDIELHIVLPWVRVGLSASTGQDKERTVHSWSTSSLTWTNVAKKENE 255
Db 244 LYSDE-----TIATLVHSDLY-----MEANLDYEFKLTSAKYT-VATARKR 285
Qy 256 NKYIT 260
Db 286 HEHIT 290

```

; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-09-977-260-4

Query Match      6.6%; Score 91; DB 9; Length 675;
Best Local Similarity 23.0%; Pred. No. 2.2;
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;

QY 72 INFESTP---YTSRIA--DGLAFFIAP-----PDSVSIYHGGL 106
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 66 VNLEQTPVERQYPFQIVYKDGLLVYASNEESRQWLKALQKEIRGNPHLLVYHSGF- 124
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 107 GLFPNANTL---NNSSTSENQTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFPDRVLKIPRAVPV-LK 181
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 210 EL--HTVLPEW--VRVGLSASTGQD-----KERNTVH-----SWFTSSLWTNVAKEN 254
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 233 QYIPREDFPDWMQVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 255 ENKY 258
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 291 LDDY 294

RESULT 12
US-09-977-269-4
; Sequence 4, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-09-977-269-4

Query Match      6.6%; Score 91; DB 10; Length 675;
Best Local Similarity 23.0%; Pred. No. 2.2;
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;

QY 72 INFESTP---YTSRIA--DGLAFFIAP-----PDSVSIYHGGL 106
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 66 VNLEQTPVERQYPFQIVYKDGLLVYASNEESRQWLKALQKEIRGNPHLLVYHSGF- 124
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 107 GLFPNANTL---NNSSTSENQTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFPDRVLKIPRAVPV-LK 181
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 210 EL--HTVLPEW--VRVGLSASTGQD-----KERNTVH-----SWFTSSLWTNVAKEN 254
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 233 QYIPREDFPDWMQVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 255 ENKY 258
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 291 LDDY 294

RESULT 11
US-09-977-260-4
; Sequence 4, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-10-186-399-3

Query Match      6.6%; Score 91; DB 9; Length 675;
Best Local Similarity 23.0%; Pred. No. 2.2;
Matches 56; Conservative 44; Mismatches 72; Indels 72; Gaps 16;

QY 72 INFESTP---YTSRIA--DGLAFFIAP-----PDSVSIYHGGL 106
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 66 VNLEQTPVERQYPFQIVYKDGLLVYASNEESRQWLKALQKEIRGNPHLLVYHSGF- 124
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 107 GLFPNANTL---NNSSTSENQTTKAASSNVAVFDTYLNPDYGD-----PNVIHIGID 158
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 125 --FVDGKFLCCQSCKAAPGCTLWEAYANLHTAVNEEKHRVPTFPDRVLKIPRAVPV-LK 181
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 159 VNSIRSKVT-AKWDWONGK-----IATAHISYNSVSKRLSVTSYYAGSKPATLSYDI 209
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 182 MDAPSSSTTLAQYDNESKKNYGSQPPSSSTSLAQYDSNSKKI-----YGSQP---NFNM 232
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 210 EL--HTVLPEW--VRVGLSASTGQD-----KERNTVH-----SWFTSSLWTNVAKEN 254
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 233 QYIPREDFPDWMQVRKLKSSSESVASSNQKERNVNHHTTSKISWEPPES--SSSEEEEN 290
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
QY 255 ENKY 258
   :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 291 LDDY 294

RESULT 10
US-10-186-399-3
; Sequence 3, Application US/10186399
; Patent No. US20020173481A1
; GENERAL INFORMATION:
; APPLICANT: Ekman, Niklas
; APPLICANT: Arighi, Elena
; APPLICANT: Vastrik, Imre
; APPLICANT: Tamagnone, Luca
; APPLICANT: Alitalo, Kari
; TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE
; FILE REFERENCE: 28113/31941A
; CURRENT APPLICATION NUMBER: US/10/186,399
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 08/320,432
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-10-186-399-3
```

[illegible]

Query Match 6.5%; Score 89.5; DB 10; Length 632;
Best Local Similarity 21.1%; Pred. No. 2.8;
Matches 58; Conservative 37; Mismatches 97; Indels 83; Gaps 15;

QY 11 FDPNQEDLIFOCHAT-STNNVLQV-----TKLDSAGNPVSSSAGRVLYSAPRLWE 60
Db 210 FDLNOKYTFIIGSATGAANNKHQIGVTLFEAYFTKPTTEANPVDIELG----- 257

QY 61 DSAVLTSFDTLIINFEISTPYTSRIADGLAFFIAPDPDSVISYHGGFLGLFPNANTLNNSST 120
Db 258 -----TAFDPLNHPEPIGLKATDEVDGDI-----KDITVEF-----NDIDTSKP 296

QY 121 SENQTTTKAAS-----SNVVAVEFDYLNPDYG-----DP-NYIHIGIDV----- 159
Db 297 GAYRVYKVVNSYGESDEKTEIEVVYTKPTITAHDTIKKDLAFDPLNVEPIGLKATDPI 356

QY 160 -NSIRSKVTAKW-DWQNGKIATAHLSYNSVS-----KRLSVTSYYAGSKPATLSYDIE 210
Db 357 DGDITDKIAVKFNVDTSKPGKYHVYTKVINSYKIDEKTEIVTVY---TKPSIVAH DVE 413

QY 211 L--HTVLP--EMVRVGLSAS---TGQDKERNVTVHS 238
Db 414 IKKDTAFDPLNVEPIGLKATDPIDGDTDKITVES 448

Search completed: February 26, 2003, 16:52:28
Job time : 26.2285 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:30 ; Search time 32.2996 Seconds
(without alignments)
1089.120 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGLSFSFTKFDPNQEDLIF.....LMTNVAKENENKVIITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

	1:	2:	3:	4:	5:	6:	7:	8:	9:	10:	11:	12:	13:	14:	15:	16:	17:	18:	19:	20:	21:	22:	23:
	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	99.4	264	20	AAW87973
2	1362	99.4	264	22	AAG62890
3	1362	99.4	286	22	AAG62894
4	800.5	58.4	234	22	AAG62901
5	797	58.2	303	22	AAG62898
6	642	46.9	290	21	AAV58736
7	637	46.5	290	21	AAV58737
8	552	40.3	237	16	AAV47465
9	463	33.8	285	20	AAV06812
10	441	32.2	285	20	AAV06814

11	440	32.1	285	20	AAV06811	R. pseudoacacia le
12	439	32.0	285	20	AAV06815	Amino acid sequenc
13	422	30.8	275	21	AAV58738	Pea lectin. Pisum
14	420	30.7	242	15	AAV45911	Bauhinia purpurea l
15	420	30.7	262	20	AAI27265	B. purpurea native
16	413	30.1	242	15	AAV45912	Bauhinia purpurea l
17	378.5	27.6	239	16	AAV47467	Pea lectin, ille.
18	366	26.7	228	16	AAV47466	Pea lectin, 2ltn.
19	345	25.2	236	16	AAV4807	PNA lectin subunit
20	321.5	23.5	274	20	AAV06813	R. pseudoacacia le
21	295	21.5	681	23	ABV93833	Herbicidally activ
22	288	21.0	715	23	ABV92849	Herbicidally activ
23	270.5	19.7	282	21	AAV25452	Herbicidally activ
24	264	19.3	265	10	AAV91967	Pinus radiata cell
25	259	18.9	711	23	ABV93337	Sequence of arceli
26	256.5	18.7	691	23	ABV93681	Herbicidally activ
27	244	17.8	632	21	AAV25109	Herbicidally activ
28	241	17.6	652	23	ABV93374	Pinus radiata cell
29	229.5	16.8	677	23	ABV92654	Herbicidally activ
30	223.5	16.3	283	23	ABV91344	Herbicidally activ
31	218.5	15.9	272	23	ABV91343	Herbicidally activ
32	217	15.8	674	23	ABV93878	Herbicidally activ
33	214.5	15.7	675	23	ABV92029	Herbicidally activ
34	213	15.5	244	10	AAV93640	Sequence of a Phas
35	208.5	15.2	669	23	ABV93142	Herbicidally activ
36	201.5	14.7	688	23	ABV93318	Herbicidally activ
37	194.5	14.2	693	23	ABV92247	Herbicidally activ
38	188.5	13.8	616	23	ABV93900	Herbicidally activ
39	188.5	13.8	649	23	ABV9134	Herbicidally activ
40	186	13.6	242	23	ABV91342	Herbicidally activ
41	185	13.5	658	23	ABV9107	Herbicidally activ
42	184.5	13.5	685	23	ABV93319	Herbicidally activ
43	181	13.2	627	23	ABV91918	Herbicidally activ
44	176	12.8	667	23	ABV92524	Herbicidally activ
45	174.5	12.7	623	23	ABV91919	Herbicidally activ

ALIGNMENTS

RESULT 1

AAW87973
ID AAW87973 standard; Protein; 264 AA.

AC AAW87973;

DT 13-APR-1999 (first entry)

A lectin derived progenitor cell preservation factor.

Lectin derived progenitor cell preservation factor; progenitor cell;
haematopoietic cell; cultured cell preservation; anticancer therapy;
myeloblastic therapy; sickle-cell anaemia; ablative therapy protection;
FLK2/FLT3 receptor.

OS Dolichos lab lab.

XX WO9859038-A1.

PD 30-DEC-1998.

PF 23-JUN-1998; 98WO-US13046.

PR 24-JUN-1997; 97US-0881189.

PA (IMCL-) IMCLONE SYSTEMS INC.

PA (HEG) UNIV CALIFORNIA.

PI Chrispeels MJ, Colucci MG, Moore JG;

DR WPI: 1999-081274/07.

XX N-PSDB; AAX03593.

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;
 KW alpha-amylase inhibitor gene.
 XX
 OS Synthetic.
 OS Unidentified.
 OS Dolichos lab lab.
 XX
 PN WO200149851-A1.
 XX
 PD 12-JUL-2001.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 XX 30-DEC-1999; 99WO-US31307.
 XX
 PA (PHYL-) PHYLOGIX LLC.
 XX
 PI Colucci MG, Chrispeels MJ, Moore JG;
 XX
 DR WPI; 2001-441882/47.
 DR N-PSDB; AAH42295.
 XX
 PT Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics
 XX
 PS Example 1; Page 59; 173pp; English.
 XX
 CC The present sequence represents fusion protein of alpha-amylase inhibitor
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting
 CC lectin). The specification describes a composition of one or more members
 CC of FRIL family of progenitor cell preservation factors. The composition
 CC is useful for alleviating or reducing the hematopoietic progenitor
 CC cell-depleting activity of a therapeutic treatment, including
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of
 CC FRIL compositions to a patient prior to treatment of the patient with
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting
 CC activity alleviates or reduces the hematopoietic progenitor
 CC cell-depleting activity of the therapeutic treatment in the patient.
 CC FRIL family members are useful for isolating population of progenitor
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is
 CC administered to reduce progenitor cell depleting effects of
 CC chemotherapeutics, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients' hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 286 AA;
 Query Match 99.4%; Score 1362; DB 22; Length 286;
 Best Local Similarity 99.6%; Pred. No. 3.5e-124;
 Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 AGSLFSFTKPNQBDLFIQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWE 60
 DB 23 AQSLFSFTKPNQBDLFIQGHATSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRLWE 82
 OY 61 DSAVLTSFTDIIINFEISTPYTRIADGLAFFIAPDPSVTSYHGGFLGLFPNANTLNSST 120
 DB 83 DSAVLTSFTDIIINFEISTPYTRIADGLAFFIAPDPSVTSYHGGFLGLFPNANTLNSST 142
 OY 121 SENQTTTAASSNVVAVPEDTYLNPDYGPDPNYTHIGIDVNSIRSKVTAKDWONGKIATA 180
 DB 143 SENQTTTAASSNVVAVPEDTYLNPDYGPDPNYTHIGIDVNSIRSKVTAKDWONGKIATA 202

OY 181 HISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHSWS 240
 DB 203 HISYNSVSKRLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHSWS 262
 OY 241 FTSSLWNTNVAKENKENKYITRGVL 264
 DB 263 FTSSLWNTNVAKENKENKYITRGVL 286
 RESULT 4
 AAG62901
 ID AAG62901 standard; Protein; 234 AA.
 XX
 AC AAG62901;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Amino acid sequence of a partial FRIL polypeptide.
 XX
 KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.
 XX
 OS Sphenostylis stenocarpa.
 XX
 PN WO200149851-A1.
 XX
 PD 12-JUL-2001.
 XX
 PF 30-DEC-1999; 99WO-US31307.
 XX
 PR 30-DEC-1999; 99WO-US31307.
 XX
 PA (PHYL-) PHYLOGIX LLC.
 XX
 PI Colucci MG, Chrispeels MJ, Moore JG;
 XX
 DR WPI; 2001-441882/47.
 XX

Legume Progenitor cell preservation factors for in vivo or ex vivo
 preservation of hematopoietic progenitor cells and as therapeutics for
 alleviating/reducing progenitor cell-depleting activity of cancer
 therapeutics
 Example 22; Page 116-117; 173pp; English.

The present sequence is a partial a FRIL (Flk2/Flt3 tyrosine kinase
 receptor-interacting lectin) protein. The specification describes a
 composition of one or more members of FRIL family of progenitor cell
 preservation factors. The composition is useful for alleviating or
 reducing the hematopoietic progenitor cell-depleting activity of a
 therapeutic treatment, including radiotherapeutic and/or
 chemotherapeutic treatments. Administration of FRIL compositions to a
 patient prior to treatment of the patient with a therapeutic treatment
 having a hematopoietic progenitor cell-depleting activity alleviates or
 reduces the hematopoietic progenitor cell-depleting activity of the
 therapeutic treatment in the patient. FRIL family members are useful for
 isolating population of progenitor cells, hemangioblasts, and mesenchymal
 stem cells. The composition is administered to reduce progenitor cell
 depleting effects of chemotherapeutics, so that the patient can receive
 a higher dose of the chemotherapeutic and preferably recover from cancer.
 It is also administered to patients having, or predisposed to developing
 a condition where the patients' hematopoietic progenitor cells are
 depleted, such as severe combined immunodeficiency or aplastic anemia.
 The isolated mesenchymal cells are useful for tissue repair.

SQ Sequence 234 AA;

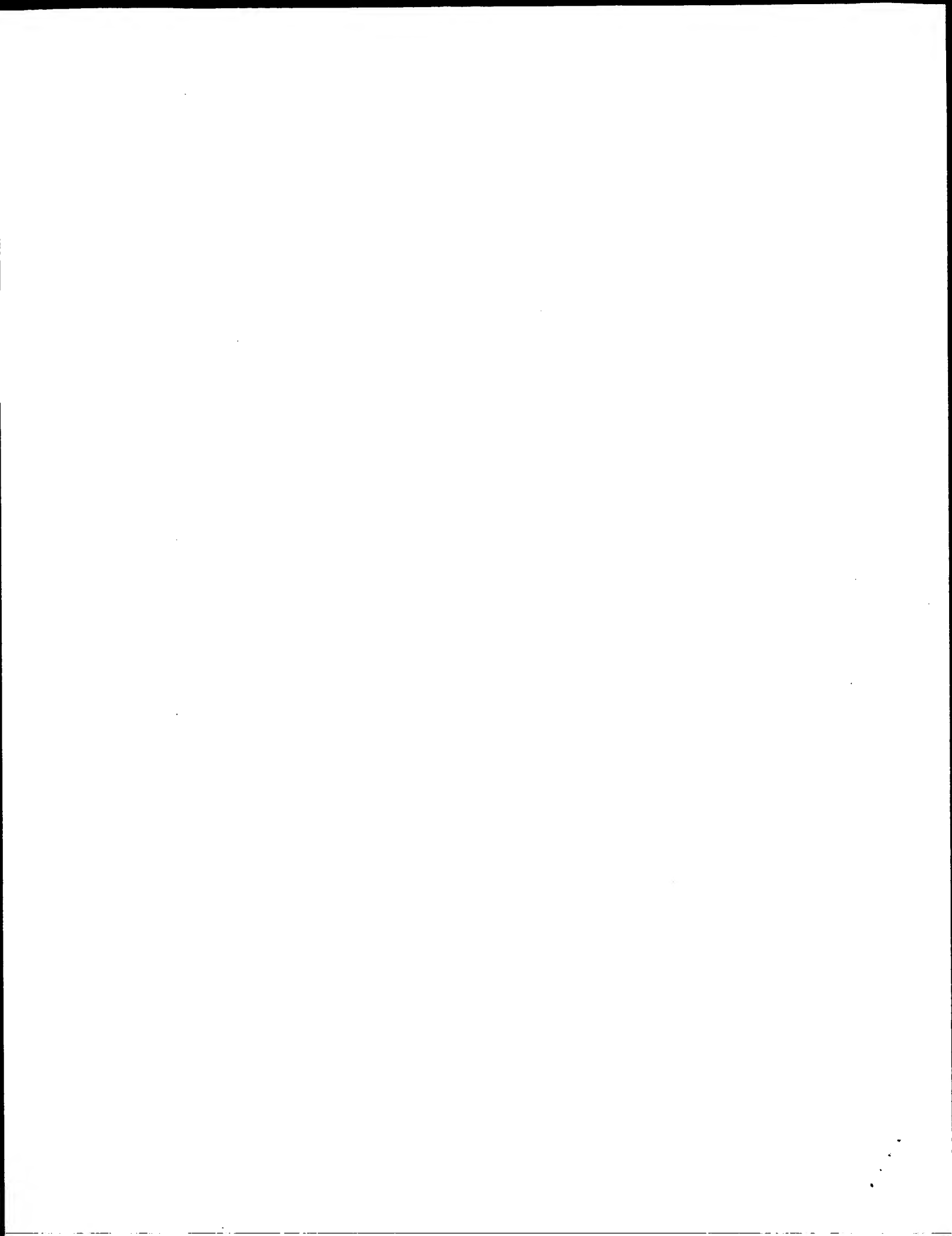
Query Match 58.4%; Score 800.5; DB 22; Length 234;
 Best Local Similarity 66.7%; Pred. No. 1.3e-69;
 Matches 156; Conservative 30; Mismatches 45; Indels 3; Gaps 2;

CC	depleted, such as severe combined immunodeficiency or aplastic anemia.
CC	The isolated mesenchymal cells are useful for tissue repair.
XX	
SQ	Sequence 303 AA;
	Query Match 58.2%; Score 797; DB 22; Length 303;
	Best Local Similarity 60.5%; Pred. No. 4e-69;
	Matches 161; Conservative 39; Mismatches 60; Indels 6; Gaps
Qy	1 AGSLSFSTFKFDPNQBDLIFQGHATSTNNVLOVTKLDGAGNPVSSAGRVLYSAPLRLWE 60 :
Dd	1 AQSLSENFTEKFDLDQDLDFIQGDATSTNNVLQT.KLDGGNPGVASGVGRVLSAFPFLWE 60 :
Qy	61 DSAVLTSFDTIINFETISPTYSRIADGLAFFIAPPDSVI--SYHGGEGLCFEPNANTLNNS 118 : : : :
Dd	61 NSNAVSSFENLIITQISTHPHYAAGFAFFLAPHDTVPPNSWGKFGLGY--SNVERNS 118 : : : :
Qy	119 STSENQTTTKA-ASSNVAVVEFTYLPNDYGDPNYTHIGIDVNSRSKVAKWDNQNGKI 177 :
Dd	119 PTSENQSGVDNTSDRVVAVEEFTTFPNAI-DPNYRHIGIDVNSIKSKETAKEWONGKT 177 :
Qy	178 ATAHISVNSVKRLSVTSYYAGSKPATLSVDIELHTVLPFWVRVGLSASTGQDKERNTVH 237 :
Dd	178 ATARISVNSAKKSSTVTTFPGMEVVALSHDVLHAELPEWRVVGSLASTGEEKOKNTII 237 :
Qy	238 SWSFTSSLWTNVAKKENKENKYITRGV 263
Dd	238 SWSFTSSLKNNEVPEDMYIANVV 263 :
RESULT 6	
ID	AAY58736 standard; Protein; 290 AA.
XX	
AC	AAY58736;
DD	25-APR-2000 (first entry)
DE	Jack bean concanavalin A lectin.
KW	Concanavalin A; Con A; lectin; jack bean; insecticide; transgenic plant; Brassica; insect resistance.
OS	Canavalia ensiformis.
FH	Key Location/Qualifiers
FT	Modified-site 152
FT	/note= "N-glycosylated"
PX	WO200001223-A1.
PD	13-JAN-2000.
PF	02-JUL-1999; 99WO-SE01209.
PR	03-JUL-1998; 98SE-0002425.
PA	(PLAN-) PLANT SCI SVERTIGE AB.
PI	Ahman I, Melander M, Vamling K;
DR	WPI; 2000-160693/14. N-PSDB; AAZ58017.
PT	Novel lectins used to produce transgenic Brassica plants which are resistant to insects -
PS	Example 2a; Fig 4; Slpp; English.
CC	The present sequence is that of the concanavalin A (Con A) lectin of jack bean. The present invention relates to: a transformed CC Brassica plant that is resistant to certain insect pests; an CC expression cassette containing DNA that codes for at least 1 CC

DR	XX	WPI; 1995-098720/13.	XX	28-AUG-1998;	98WO-GB02612.
XX	PT	New cpds. inhibiting binding of serum amyloid P to amyloid	XX	29-AUG-1997;	97GB-0018413.
PT	PT	fibrils - produced by computer assisted molecular design, useful	XX	(ALIZ-) ALIZYME THERAPEUTICS LTD.	
PT	PT	for preventing, treating or diagnosing amyloidosis, e.g.	XX	Bardocz 2M, Fish NW, Palmer RMJ, Pusztai AJ;	
XX	XX	Alzheimer's disease	XX	WPI; 1999-228935/19.	
PS	PS	Claim 7; Fig 4; 72pp; English.	XX	Use of Robina pseudoacacia lectin in medicine	
XX	CC	The sequences given in AAR74763-70 represent various pentraxins and	XX	Disclosure; Fig 1; 37pp; English.	
CC	CC	legume lectins. All these proteins show structural homology and may	XX	The invention relates to Robinia pseudoacacia lectins which can be used	
CC	CC	be used in the molecular design of a molecule for the inhibition of	CC	in medicine. The lectin is used for the control of mucosal cell	
CC	CC	serum amyloid P (SAP) binding to amyloid fibrils. The similarities	CC	proliferation, for the reduction and/or treatment of damage caused by	
CC	CC	in the amino acid sequences of SAP, human and Limulus C-reactive	CC	cell damaging agents, especially in the treatment of cancer, and/or for	
CC	CC	protein (CRP) and female hamster protein suggests that they may have	CC	the reduction and/or treatment of metabolic disorders. It is especially	
CC	CC	similar 3-D structures. The jellyroll topology of the pentraxins is	CC	useful for the treatment of mucositis in mammalian cells and/or tissues,	
CC	CC	reminiscent of the picornavirus coat proteins which also have	CC	particularly human cells and/or tissues (especially mucous cells	
CC	CC	pentameric structures. However, pentaxins resemble more closely	CC	including mucous membrane). Use of the lectin is effective and does not	
CC	CC	legume lectins such as concanavaline A and pea lectin. Alignment	CC	damage the gut as is the case with cytotoxic drugs and radiation. The	
CC	CC	of sequences on the basis of topologically equivalent features of the	CC	present sequence represents a polypeptide subunit of R. pseudoacacia	
CC	CC	three dimensional structures shows that helices occupy different	XX	lectin RPBAl.	
CC	CC	positions in the pentraxins and legume lectins and that the amino acid	XX	Sequence 285 AA;	
CC	CC	sequences of the two families have identities of only approx. 11%. The	Query Match	33.8%; Score 463; DB 20; Length 285;	
CC	CC	two main helices in SAP occur before and after strand L, whereas the	Best Local Similarity	43.9%; Pred. No. 1.2e-36;	
CC	CC	helices in the legume lectins occur at the C-terminus of strand J.	Matches 115; Conservative	40; Mismatches 81; Indels 26; Gaps 10;	
CC	CC	There is a long insertion between the end of the helix after strand D			
CC	CC	and the beginning of strand E in the lectins relative to pentraxins.			
CC	CC	Strands G, H and I together with the type IV beta-hairpin between H and			
CC	CC	I are identical in both SAP and pea lectin. The so-called pentraxin			
CC	CC	octapeptide signature sequence, HXCXS/TWXS, is in this region so it is			
CC	CC	not conserved in the legume lectins.			
XX	XX	Sequence 237 AA;			
XX	XX	Query Match			
XX	XX	Best Local Similarity			
XX	XX	Matches 115; Conservative			
XX	XX	Score 552; DB 16; Length 237;			
XX	XX	Pred. No. 2e-45;			
XX	XX	Mismatches 61; Indels 20; Gaps 3;			
QY	25	TSTNNLVQVTKLDSAGNPVSSAGRLVYSAPLRLEDSAVLTSTFTIINFESTPYTSRI 84	QY	2	GSLSFSFTKEDPNOEDLIFQCHATSTN-NVLQVTKLDSAGNPVSSAGRLVYSAPLRLEW 60
Db	2	TGTDGNLELTRVSSNGSEGSVGRALFAPVHIWESSAATVSFEATFAFLIKSP-DSHP 60	Db	33	GSLSFSFPKFKSQPDLITFOSDALVTSGVLQLTVND--GRVYDSIGRLVYAAPFQIWD 90
QY	85	ADGLAFFIAPDPSVI--SYHGFGFLGFPNANTLNNSSTSENQTTTKAASNNVAVFEEDY 142	QY	61	D-SAVLTSTFTIINFESTPYTSRIADGLAFFIAPDPSVISYH-GGFLGLFPNANTLNNS 118
Db	61	ADGLAFFISNIDSSIPSGSTGILGLFPDAN-----ADTIVAVELDTY 103	Db	91	STTGNAVASFYTSFSFIKAPNEGKTADGLVFFLAPVSGTQPLKGGLLGLF----- 141
QY	143	LNPDYGDPNYTHIGIDVNSIRSKVTAKWDQNGKIATATAHISYNSVSKRLSVTSYYAGSKP 202	QY	119	STSENQTTTKAASNNVAVFEEDYINPDYGDPNYTHIGIDVNSIRSKVTAKWDQNGKIA 178
Db	104	PNTDGDPSYPHIGIDIKSVRSKKTAKWNQDKVGTCHAILIYNSVDKRLSAVSYPNADA 163	Db	142	--KDESYNK--SNQIYAVEDETFERNVAV-DPNGIHMGIDVNSIQSVRTVRWDWANGEVA 195
QY	203	ATLSYDIELHTVLPWVRVGLSASTGQDKERNVTWSFTSSLWN 248	QY	179	TAHISYNSVSKRLSVTSYYAG-SKPATLSYDIELHTVLPWVRVGLSASTGQDK---ERN 234
Db	164	TSVSYDVLNDVLPWVRVGLSASTGLYKETNTIISWFTSKLSN 209	Db	196	NVFISYEASTKSLTASLVYPSLEKSFILSAIVDLKKVLPWVRVGTATTGLSEYVQTN 255
RESULT 9			QY	235	TVHSWSFTSSL--WTNVAKKEN 254
RAY06812			Db	256	DVLSWSFESNLPGGSNVASVKN 277,
ID	AAV06812	standard; protein; 285 AA.	RESULT 10		
XX	XX	AAV06812	AAV06814		
XX	XX	AAV06812;	ID	AAV06814	standard; protein; 285 AA.
XX	XX	23-JUN-1999 (first entry)	XX	XX	AAV06814;
XX	XX	R. pseudoacacia lectin RPBAl polypeptide subunit b.	XX	XX	23-JUN-1999 (first entry)
XX	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;	XX	XX	Amino acid sequence of lectin RPBAl from R. pseudoacacia seed.
XX	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;	XX	XX	Robinia pseudoacacia; lectin; medicine; mucosal cell proliferation;
XX	XX	caner; metabolic disorder; mucositis; cytotoxic.	XX	XX	caner; metabolic disorder; mucositis; cytotoxic.
XX	XX	Robinia pseudoacacia.	XX	XX	Robinia pseudoacacia.
XX	XX	WO9911278-A1.	XX	XX	WO9911278-A1.
XX	XX	11-MAR-1999.	XX	XX	11-MAR-1999.
PD	PD		XX	XX	

XX 29-AUG-1997; 97GB-0018413.
 XX (ALIZ-) ALIZYME THERAPEUTICS LTD.
 XX Bardocz ZM, Fish NW, Palmer RMJ, Pusztai AJ;
 XX WPI; 1999-228935/19.
 XX Use of Robina pseudoacacia lectin in medicine
 XX Disclosure; Fig 2; 37pp; English.
 XX The invention relates to Robinia pseudoacacia lectins which can be used
 XX in medicine. The lectin is used for the control of mucosal cell
 XX proliferation, for the reduction and/or treatment of damage caused by
 XX cell damaging agents, especially in the treatment of cancer, and/or for
 XX the reduction and/or treatment of metabolic disorders. It is especially
 XX useful for the treatment of mucositis in mammalian cells and/or tissues,
 XX particularly human cells and/or tissues (especially mucous cells
 XX including mucous membrane). Use of the lectin is effective and does not
 XX damage the gut as is the case with cytotoxic drugs and radiation. The
 XX present sequence represents a R. pseudoacacia lectin RpbAII. This lectin
 XX is composed of the subunit- polypeptide c (AA06813).
 XX Sequence 285 AA;
 SQ Query Match 32.0%; Score 439; DB 20; Length 285;
 Best Local Similarity 42.5%; Pred. No. 2.7e-34;
 Matches 111; Conservative 39; Mismatches 87; Indels 24; Gaps 9;
 QY 2 GSLGSEFTKDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSAGRVLYSAPLRLWE 60
 DB 33 GSLGSEFTKDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSAGRVLYSAPLRLWE 60
 QY 61 D-SAVLTSFDTIINFEISTPTSTRADGLAFPIAPDPSVISYHGGFLGFLFPNANTLNSS 119
 DB 92 STTGNAVSEFTVSFIIQAPNPATTADGLAFPIAPDPSVISYHGGFLGFLFPNANTLNSS 119
 QY 120 TSENQTTKAASSNVVAFEDFTYLNPDYDGNVTHIGIDVNSIRSKVTAKDWONGKIAI 179
 DB 149'-----SNQIVAFEDFTYLNPDYDGNVTHIGIDVNSIRSKVTAKDWONGKIAI 179
 QY 180 AHISYNSVKRLSVTSYAGSKPATLSYDI-ELHTVLPWVRVGLSASTGQDK---ERNT 235
 DB 197 VFISYEASTKSLTASIVPSLETSTFIHAIVDKVLPWVRVGLSASTGQDK---ERNT 235
 QY 236 VHSWFTSSL--WTNVAKEN 254
 DB 257 VLWSFESNLPDGNVSVAKN 277
 RESULT 13
 AAY58738
 ID AAY58738 standard; Protein; 275 AA.
 XX AC AAY58738;
 XX 25-APR-2000 (first entry)
 XX Pea lectin.
 XX Lectin; pea; insecticide; transgenic plant; Brassica;
 XX insect resistance.
 XX Pisum sativum.
 XX WO200001223-A1.
 XX 13-JAN-2000.
 XX 02-JUL-1999; 99WO-SE01209.
 XX

PR 03-JUL-1998; 98SE-0002425.
 XX (PLAN-) PLANT SCI SVERIGE AB.
 XX Ahman I, Melander M, Vamling K;
 XX WPI; 2000-160693/14.
 XX N-PSDB; AA258019.
 XX Novel lectins used to produce transgenic Brassica plants which are
 XX resistant to insects -
 XX Example 2b; Fig 9; 51pp; English.
 XX The present sequence is that of pea cv. Lincoln lectin. The
 XX invention relates to: a transformed Brassica plant that is
 XX resistant to certain insect pests; an expression cassette
 XX containing DNA that codes for at least 1 lectin selected from Con A
 XX (see AAY58736), modified Con A (see AAY58737) and pea lectin; a new
 XX transgenic plant cells containing at least 1 copy of the DNA; a new
 XX lectin derived from the Jack bean Con A gene; a method of imparting
 XX resistance to insects selected from blossom beetles (pollen
 XX beetles) of the genus Meligethes, flea beetles of the genus
 XX Phyllotreta, and root flies of the genus Delia; and a method for
 XX protecting a plant against infestation by insects of these genera.
 XX Sequence 275 AA;
 SQ Query Match 30.8%; Score 422; DB 21; Length 275;
 Best Local Similarity 39.3%; Pred. No. 1.2e-32;
 Matches 101; Conservative 41; Mismatches 87; Indels 28; Gaps 6;
 QY 5 SFSFTKDPNQEDLIFQGHATSTNNVLQVTKLDSAGNPVSSAGRVLYSAPLRLWE-DSA 63
 DB 35 SFLITKESPPQNLIFQGDGTYTTRKLT/LTK-----AVKNTVGRALYSSPFIWDRETG 88
 QY 64 VLTSFDTIINFEISTPTSTRADGLAFPIAPDPSVISYHGGFLGFLFPNANTLNSSSTEN 123
 DB 89 NVANFVTSFTVINAPNSYNVADGFTFIAPDPSVISYHGGFLGFLFPNANTLNSSSTEN 141
 QY 124 QTTTKAASSNVVAFEDFTYLNPDYDGNVTHIGIDVNSIRSKVTAKDWONGKIAI 182
 DB 142 TTQT-----VAVEEDFTYNAAMDPSNRDRHIGIDVNSIXSVNTKSKWLONGEEANVVI 194
 QY 183 SYNSVSKRLSVTSYAGS-----KPATLSYDIELHTVLPWVRVGLSASTGQDKERNIV 236
 DB 195 AFNAATNVLTVSLTYPNLSLEENVTSTYLSDVSLKDVVPEWVRIGFSATTGAETAAHEV 254
 QY 237 HSWFTSSLTNVAKKE 253
 DB 255 LWSFHSLSLCTSSSKQ 271
 RESULT 14
 AAR45911
 ID AAR45911 standard; protein; 242 AA.
 XX AC AAR45911;
 XX 18-JUL-1994 (first entry)
 XX Bauhinia purpurea lectin larvicidal deriv.
 XX Non-lysine; insect larvae; sorghum; wheat; oat; rye; rice; European;
 XX corn borer.
 XX Bauhinia purpurea.
 XX WO9402514-A.
 XX 03-FEB-1994.
 XX 22-JUL-1993; 93WO-US06946.
 XX



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:43:23 : Search time 25.0487 Seconds
(without alignments)
2171.628 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLSFSTKFDPNQEDLIF.....LWTVAKKENKIYTRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum Match 0%

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteriap:.*
17: sp_archeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362	99.4	272	10 Q9ZTA9	Q9ZTA9 dolichos la
2	797	58.2	279	10 Q9M7M4	Q9M7M4 phaseolus v
3	642	46.9	290	10 Q947H0	Q947H0 canavalia e
4	641	46.8	290	10 O04672	O04672 canavalia b
5	631.5	46.1	293	10 P93537	P93537 sophora jap
6	583.5	42.6	284	10 Q9FVU9	Q9FVU9 sophora fla
7	576.5	42.1	266	10 P93536	P93536 sophora jap
8	502	36.6	286	10 P93248	P93248 maackia amu
9	496	36.2	254	10 Q43376	Q43376 arachis hyp
10	494	36.1	282	10 P93247	P93247 maackia amu
11	493	36.0	280	10 Q43374	Q43374 arachis hyp
12	486	35.5	254	10 Q43377	Q43377 arachis hyp
13	459.5	33.5	258	10 Q9FVF8	Q9FVF8 ulex europe
14	459	33.5	256	10 P93246	P93246 maackia amu
15	454.5	33.2	273	10 Q8RVY4	Q8RVY4 phaseolus c
16	452	33.0	285	10 Q9ZWP6	Q9ZWP6 robinia pse

17	443.5	32.4	247	10 Q8WLR6	Q8WLR6 griffonia s
18	436	31.8	251	10 Q93X49	Q93X49 lens culina
19	436	31.8	251	10 Q8VXF2	Q8VXF2 lens culina
20	435	31.8	251	10 Q93X50	Q93X50 lens culina
21	434	31.7	251	10 Q93X41	Q93X41 lens odemen
22	434	31.7	251	10 Q93WH6	Q93WH6 lens culina
23	434	31.7	251	10 Q8W4Y4	Q8W4Y4 lens lamott
24	433.5	31.6	273	10 Q8RVH2	Q8RVH2 phaseolus v
25	432.5	31.6	275	10 Q8RVH1	Q8RVH1 phaseolus v
26	432	31.5	279	10 Q49899	Q49899 medicago sa
27	431.5	31.5	251	10 Q8WLR7	Q8WLR7 griffonia s
28	431.5	31.5	275	10 Q8RVX5	Q8RVX5 phaseolus v
29	430.5	31.4	275	10 Q8RW23	Q8RW23 phaseolus c
30	425	31.0	251	10 Q93X48	Q93X48 lens ervoid
31	425	31.0	274	10 Q43628	Q43628 phaseolus v
32	423.5	30.9	268	10 P93458	P93458 phaseolus l
33	423.5	30.9	278	10 Q9LEB8	Q9LEB8 phaseolus l
34	420.5	30.7	275	10 Q8RVY1	Q8RVY1 phaseolus v
35	420.5	30.7	276	10 Q40750	Q40750 phaseolus a
36	420	30.7	251	10 Q93X42	Q93X42 lens nigric
37	419.5	30.6	278	10 Q9LEP9	Q9LEP9 phaseolus l
38	416.5	30.4	255	10 Q40912	Q40912 phaseolus l
39	415.5	30.3	273	10 Q8RVX9	Q8RVX9 phaseolus v
40	415	30.3	251	10 Q8RW33	Q8RW33 lathyrus sa
41	412.5	30.1	275	10 Q8RVX6	Q8RVX6 phaseolus v
42	411.5	30.0	273	10 Q42411	Q42411 medicago sa
43	411.5	30.0	275	10 Q8RVH3	Q8RVH3 phaseolus v
44	410.5	30.0	275	10 Q8RVV0	Q8RVV0 phaseolus v
45	409	29.9	285	10 Q9ZWP5	Q9ZWP5 robinia pse

ALIGNMENTS

RESULT 1

Q9ZTA9 PRELIMINARY; PRT; 272 AA.
ID Q9ZTA9
AC Q9ZTA9
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mannose lectin.
GN FRIL.
OS Dolichos lab lab (Field bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=35936;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COTYLEDON;
RX MEDLINE=99110944; PubMed=9892687;
RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;
RT "CDNA cloning of FRIL, a lectin from Dolichos lablab--that preserves
hematopoietic progenitors in--suspension culture";
RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).
DR HSSP; AF067417; AAD10734.1; -
DR EMBL; P02866; IONA
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGME_BETA; UNKNOWN_1.
KW Lectin.
SQ SEQUENCE 272 AA; 29900 MW; EA6C004307441495 CRC64;

Query Match 99.4%; Score 1362; DB 10; Length 272;
Best Local Similarity 99.6%; Pred. No. 3.4e-96;
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AGSLSFSTFKFDPNQEDLIFQGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWE	60
Db	9	AQSLSFSTFKFDPNQEDLIFQGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWE	68
QY	61	DSAVLTSFDTIINFESTPYTSRIADGLAFFIAPDPSVISYHGGFLGLFPNANTLNNSST	120
Db	69	DSAVLTSFDTIINFESTPYTSRIADGLAFFIAPDPSVISYHGGFLGLFPNANTLNNSST	128
QY	121	SENQTTTKAASNNAVVEFDYTLNPDYGDPNYIHGIDVNSIRSKVTAKWDQWQNKIATA	180
Db	129	SENQTTTKAASNNAVVEFDYTLNPDYGDPNYIHGIDVNSIRSKVTAKWDQWQNKIATA	188
QY	181	HISYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHWS	240
Db	189	HISYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVHWS	248
QY	241	FTSSLTWNVAKKENENKYITRGVL	264
Db	249	FTSSLTWNVAKKENENKYITRGVL	272
RESULT 2			
Q9M7M4		PRELIMINARY; PRT; 279 AA.	
ID	Q9M7M4		
AC	Q9M7M4		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Mannose lectin FRIL (Fragment)		
OS	Phaseolus vulgaris (kidney bean) (French bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.		
OX	NCBI_TaxID=3885;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,		
RA	Feldman M.;		
RT	"A new lectin in red kidney bean called PvFRIL stimulates		
RT	proliferation of NIH3T3 cells expressing the Flt3 receptor.";		
RL	Submitted (JAN-1999) to the EMBL/GenBank/DBD databases.		
DR	EMBL; AF121458; AAF28739.1; -		
DR	HSP; P02866; IONA.		
DR	InterPro; IPR000985; Lectin_legA.		
DR	InterPro; IPR001220; Lectin_legB.		
DR	Pfam; PF00138; lectin_legA; 1.		
DR	Pfam; PF00139; lectin_legB; 1.		
DR	ProDom; PD000671; lectin_legA; 1.		
DR	ProDom; PD000711; lectin_legB; 1.		
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.		
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.		
KW	Lectin.		
FT	NON-TER		
SQ	SEQUENCE 279 AA; 31102 MW; F8919CF8B3EE4652 CRC64;		
Query Match	58.2%; Score 797; DB 10; Length 279;		
Best Local Similarity	60.5%; Pred. No. 4.4e-53;		
Matches	161; Conservative 39; Mismatches 60; Indels 6; Gaps		
QY	1	AGSLSFSTFKFDPNQEDLIFQGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWE	60
Db	1	AQSLSFSTFKFDPNQEDLIFQGHATSTNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWE	68
QY	61	DSAVLTSFDTIINFESTPYTSRIADGLAFFIAPDPSVI--SYHGGFTGLFPNANTLNNS	111
Db	61	NSMAYSSSEFNTLIQISTPHPYAAGDGAFFLPHDTVIPPNWSGKFLGLY--SNVFRNS	111
QY	119	STSENQTTTKA-ASSNVAVEFDYTLNPDYGDPNYIHGIDVNSIRSKVTAKWDQWQNKI	17
Db	119	PTSENQSGDVAQDNRVAAVEEDTFPNMI--DPNVRHIGIDVNSIRSKETARWQWQNK	17
QY	178	ATAHLSYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGDKERNTVH	23

Db	178	ATARISYNSASKSKSTVTTTYPGMEVVVALSHDVLHAELPEWVRVGLSASTGEKOKNII	233
QY	238	SWSFSSLTNTNNAKKENENKNIYTRGV	263
Db	238	SWSFSSLTNNKEVKPKEDMYIANVV	263
RESULT 3			
Q947H0		PRELIMINARY;	
ID	Q947H0	PRT;	290 AA.
AC	Q947H0;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Concanavalin A.		
GN	CONA.		
OS	Canavalia ensiformis (Jack bean) (Horse bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid1 I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.		
OX	NCBI_TaxID=3823;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV. 0-02;		
RA	Ranis C. Gomord V.;		
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF308777; AAL09432.1; -		
DR	InterPro; IPR000985; Lectin_legA.		
DR	InterPro; IPR001220; Lectin_legB.		
DR	Pfam; PF00138; lectin_legA; 1.		
DR	Pfam; PF00139; lectin_legB; 1.		
DR	ProDom; PD000671; Lectin_legA; 1.		
DR	ProDom; PD000711; Lectin_legB; 1.		
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN.1.		
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN.1.		
SQ	SEQUENCE 290 AA; 31480 MW; 0F2F7DBBCF547E42 CRC64;		
Query Match	46.9%;	Score 642;	DB 10; Length 290;
Best Local Similarity	53.8%;	Pred. No. 3e-41;	
Matches 134;	Conservative	Mismatches 74;	Indels 6; Gaps
QY	3	SLSFTFFDQEDLIFOGHATS--TNNLVOTKLSAGNPVSSSAGRVLYSAPRLRWED 61	
Db	36	ALHFMFNFGSKDKDLILGGDATTGTDGNLETRVSSNGSPQSGSVGRALFYAPVHIWES 95	
QY	62	SAVLRSFTIINFELSTPYTSRIADGAFFIAPPDSVI--SYHGGFGGLPPNANTLNSS 119	
Db	96	SAVVASFEATFTFLKSP-DSHPADGIAFFISNIDSSPSGSGTGRLLGLEFPDANVRNST 159	
QY	120	TSENQTTKAASNNVAVFETIYNPDYGDNPYTHIGIDVNSIRSKYTKAKWDWONGKAT 179	
Db	155	TIDFNAAYNA--DTIVAVELDTPNTDGDPSYPHIGIDIKSVRSKKTKAKNMQNGKVG 219	
QY	180	AHISYNSYKRLSVTSYYAGSKPATLSYDIEHTLTPWVRVGLSASTGQDKERTNVHSW 239	
Db	213	AHLIYNSVDKRLSAVVSYPNADSAIVSYDVLNDVLPWVRVGLSASTGLYKETNIIISW 279	
QY	240	SFTSSLTWN 248	
Db	273	SFTSKLKS 281	
RESULT 4			
O04672		PRELIMINARY;	
ID	O04672	PRT;	290 AA.
AC	O04672;		
DT	01-JUL-1997 (TREMBLrel. 04, Created)		
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Lectin (Fragment).		
OS	Canavalia brasiliensis (Brazilian jack bean).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; fabaceae; Papilionoideae; Sophoreae; Maackia.
OX NCBI_TaxID=37501;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 29-48.
RC TISSUE=BARK;
RX MEDLINE=97390228; PubMed=9249142;
RA van Damme E.J.M., Van Leuven F., Peumans W.J.;
RT "Isolation, characterization and molecular cloning of the bark lectins
from Maackia amurensis";
RL Glycoconj. J. 14:449-456(1997).
CC -1- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
CC -1- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR EMBL; U65010; AAB39934.1; .
DR HSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_lega.
DR InterPro; IPR001220; Lectin_legb.
DR Pfam; PF00138; lectin_lega; 1.
DR Pfam; PF00139; lectin_legb; 1.
DR ProDom; PD000671; Lectin_lega; 1.
DR ProDom; PD000711; Lectin_legb; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
FT NON_TER 1 1
FT SIGNAL <1 28
FT CHAIN 29 286 BARK LEUCOAGGLUTININ II.
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 207 207 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 286 AA; 31208 MW; 0F2E27617A0F6D00 CRC64;

Query Match 36.6%; Score 502; DB 10; Length 286;
Best Local Similarity 43.4%; Pred. No. 1.4e-30;
Matches 116; Conservative 42; Mismatches 85; Indels 24; Gaps 9;

QY 4 LSFSTKFPDNOEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWED- 61
DB 32 LSFSTKFPDNOEDLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWED- 61
QY 62 SAVLTSFTDIINFEISTPYTSRIAAGLAFPIAPPDSVI--SYHGGFLGLFPNANTLNNS 119
DB 91 TGSVASFTSTFTVVKAPNTITSDGLAFLAPPDSQIPSGRVSKYLGLFNNSN-----S 145
QY 120 TSENQTTTKAASSNVAVVEFDYLPNDYG--DPNYTHIGIDVNSIRSKYAKWDWONGKI 177
DB 146 DSSNQ-----IVAVEFDYFSGSDPDPNRRHIGIDVNGIESIKTVQWDWINGGV 196
QY 178 ATAHISYNSVSKRLSVTSYAGSKPA-TLSYDIELHTVLPBWVRVGLSASTG--QDKERN 234
DB 197 AFATITYLAPNKLTLASLYPSNQTSFVVAASVDLKEILPEWVRVGSAAATGYPTQVETH 256
QY 235 TVHWSFTSTSLTNVAKKENENKITYR 261
DB 257 DVLWSFTSTLEANSDAENNVHAR 283

RESULT 9
Q43376 PRELIMINARY; PRT; 254 AA.
AC Q43376;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE Mannose/glucose-binding lectin precursor (Fragment).
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. SELLIE; TISSUE=SEED;
RA Law I.J.;
RT "Cloning and expression of cDNA for mannose-binding lectin from
peanut";
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.;
DR EMBL; U22472; AAA74575.1; .
DR HSP; P02867; 2BQP.
DR InterPro; IPR000985; Lectin_lega.
DR InterPro; IPR001220; Lectin_legb.
DR Pfam; PF00138; lectin_lega; 1.
DR Pfam; PF00139; lectin_legb; 1.
DR ProDom; PD000671; Lectin_lega; 1.
DR ProDom; PD000711; Lectin_legb; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT NON_TER 1 1
FT SIGNAL <1 2
FT CHAIN 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;
SQ SEQUENCE 254 AA; 28202 MW; 3AE1EALF90BICA03 CRC64;

Query Match 36.2%; Score 496; DB 10; Length 254;
Best Local Similarity 45.6%; Pred. No. 3.5e-30;
Matches 114; Conservative 41; Mismatches 75; Indels 20; Gaps 7;

QY 3 SLFSFTKFPDNOE-DLIFOGHAT-STNNVLOVTKLSAGNPVSSAGRVLYSAPRLWE 60
DB 3 SLFSYKFKQDDERNLILQGDATFSASKGQLTKVDANGTPAKSTVGRVLHSTQVRLWE 62
QY 61 DSA-VLTSFTDIINFEISTPYTSRIAAGLAFPIAPPDSVIVH--GGFLGLFPNANTLN 117
DB 63 KSTNRLTNFQAQSFVVKSP-NDIGADGIAFFIAAPDSQIPKNSAGTGLGF----- 113
QY 118 SSTSENQTTTKAASSNVAVVEFDYLPNDYG--DPNYTHIGIDVNSIRSKYAKWDWONG 175
DB 114 ----DPQTAQNPSANQVLAFFEDTFYAGDSNGWDPNYQHIGIDVNSIKSAATKWRNG 169
QY 176 KIATAHISYNSVSKRLSVTSYAGSKPATLSYDIELHTVLPBWVRVGLSASTGQDKERN 235
DB 170 QTLNVLYTDANSKNLQVTASYPGQRYQVSYVDRLDHLPEWGRVGSASSGQQVQSHE 229
QY 236 VHSWSFTSSL 245
DB 230 LQSWSTSNL 239

RESULT 10
P93247 PRELIMINARY; PRT; 282 AA.
ID P93247
AC P93247
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE BARK LEUCOAGGLUTININ I precursor (MALBI) (Fragment).
GN LECMALBI.
OS Maackia amurensis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; fabaceae; Papilionoideae; Sophoreae; Maackia.
OX NCBI_TaxID=37501;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=BARK;
RX MEDLINE=97390228; PubMed=9249142;
RA van Damme E.J.M., Van Leuven F., Peumans W.J.;

RT "Isolation, characterization and molecular cloning of the bark lectins
 RL from *Maackia amurensis*.";
 CC Glycoconj. J. 14:449-456(1997).
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).
 CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.
 CC -!- PTM: ONLY THREE OF THE PUTATIVE GLYCOSYLATION SITES ARE OCCUPIED.
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 DR EMBL: U65009; AAB39933.1; .
 DR HSSP: P19588; LLU.
 DR InterPro: IPR000985; Lectin_legA.
 DR InterPro: IPR001220; Lectin_legB.
 DR InterPro: IPR00531; TonB_boxC.
 DR Pfam: PF00138; lectin_legA; 1.
 DR Pfam: PF00139; lectin_legB; 1.
 DR ProDom: PD000671; Lectin_legA; 1.
 DR ProDom: PD000711; Lectin_legB; 1.
 DR ProSite: PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR ProSite: PS00307; LECTIN_LEGUME_BETA; UNKNOWN.1.
 DR ProSite: PS00430; TONB_DEPENDENT_REC.1; UNKNOWN.1.
 KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family; Signal.
 FT NON_TER 1
 FT SIGNAL <1 24 BY SIMILARITY.
 FT CHAIN 25 282 BARK LEUCOAGGLUTININ I.
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 282 AA; 30879 MW; 45B3F714E1D8957A CRC64;
 Query Match 36.1%; Score 494; DB 10; Length 282;
 Best Local Similarity 43.1%; Pred. No. 5.7e-30;
 Matches 115; Conservative 42; Mismatches 86; Indels 24; Gaps 8;
 QY 4 LSFSTKFPDNEEDLFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLSAPLRLWE- 61
 DB 28 LSFSTNNFLPNEADLLFOGEASVSSTGVLQTLRVEN-GQPKYSVGRALYAAPVRIWDNT 86
 QY 62 SAVLSFSDTIINFEISTPTYSRTADGLAFIAPDPSVISYHG--GELGLFPNANTLNSS 119
 DB 87 TGSVASFSTFTFVVKAPNPSTSNGLAFIAPDPSQIPGTSGVTKVLGLFNN----- 138
 QY 120 TSENQTTTKAASNNVAVFEDTVLNPDYG--DPNYTHIGIDVNSIRSKVTAKWDWQNGKI 177
 DB 139 -----TSDSSNQIVAVFEDTVLHKNYPWDPNYRHIGIDVNGIDSIKTVQWDWINGGV 192
 QY 178 ATAHISYNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPWVRVGLSASTGQDKERNT- 235
 DB 193 APATITLAPSKTLIASLYPSNQTSFIVAASVDLKEILPEWVRVGFSAATGYPTVEVETH 252
 QY 236 -VHSNSETSSLTWNVAKKENKYLIR 261
 DB 253 DVLVSFSTLEANSDAATENNVHIAIR 279
 RESULT 11
 Q43374 PRELIMINARY; PRT; 280 AA.
 AC Q43374;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Mannose/glucose-binding lectin precursor.
 GN LEC.
 OS *Arachis hypogaea* (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC *Arachis*.

OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SELLIE; TISSUE=NODULE;
 RA Law I.J.;
 RT "Cloning and expression of cDNA for mannose-binding lectin from
 peanut.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U22469; AAA74572.1; .
 DR HSSP: P02866; LDO2.
 DR InterPro: IPR000985; Lectin_legA.
 DR InterPro: IPR001220; Lectin_legB.
 DR Pfam: PF00138; lectin_legA; 1.
 DR Pfam: PF00139; lectin_legB; 1.
 DR ProDom: PD000671; Lectin_legA; 1.
 DR ProDom: PD000711; Lectin_legB; 1.
 DR ProSite: PS00307; LECTIN_LEGUME_BETA; UNKNOWN.1.
 KW Lectin; Signal.
 FT SIGNAL 1 26
 FT CHAIN 27 280
 SQ SEQUENCE 280 AA; 31012 MW; C15B39B32F455BD5 CRC64;
 Query Match 36.0%; Score 493; DB 10; Length 280;
 Best Local Similarity 45.2%; Pred. No. 6.7e-30;
 Matches 113; Conservative 44; Mismatches 73; Indels 20; Gaps 8;
 QY 3 SLFSFTKFPDNOE-DLIFOGHAT-STNNVLQVTKLDSAGNPVSSSAGRVLSAPLRLWE 60
 DB 29 SLFSYNNFEQDDERNLILQGDATFSASKGIQLTKVDDNGTFAKSTVGRVLHSTQVRLWE 88
 QY 61 DSA-VLTSFSDTIINFEISTPTYSRTADGLAFIAPDPSVISYH--GGFLGLFPNANTLN 117
 DB 89 KSTNRLTNFQAQFSVFKSPIDNG-ADGIAPFAAPDSIPKNSAGGTGLGF-----D 140
 QY 118 SSTSENQTTTKAASNNVAVFEDTVLNPDYG--DPNYTHIGIDVNSIRSKVTAKWDWQNG 175
 DB 141 PSTAQN-----PSANQVLAUFEDTVFAQDNGWDPNYQHIGIDVNSIKSAATTKWERRNG 195
 QY 176 KIATAHISYNSVKRLSVTSYAGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERNT 235
 DB 196 QTLNVLVSYDANSRNLQVTASYPDQGRYQVSYNVLDRLYLPEWGSVGFSAASGQQYQSHE 255
 QY 236 VHSWSFTSSL 245
 DB 256 IQSWSTSTL 265
 RESULT 12
 Q43377 PRELIMINARY; PRT; 254 AA.
 AC Q43377;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Mannose/glucose-binding lectin precursor (Fragment).
 GN LEC.
 OS *Arachis hypogaea* (Peanut).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
 OC *Arachis*.
 OX NCBI_TaxID=3818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. SELLIE; TISSUE=SEED;
 RA Law I.J.;
 RT "Cloning and expression of cDNA for mannose/glucose-binding lectin
 from peanut.";
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U22473; AAA74576.1; .
 DR HSSP: P02867; 2BOP.
 DR InterPro: IPR000985; Lectin_legA.
 DR InterPro: IPR001220; Lectin_legB.

DR	Pfam:	PF00138;	lectin_legA; 1.
DR	Pfam:	PF00139;	lectin_legB; 1.
DR	ProDom:	PD000671;	Lectin_legA; 1.
DR	ProDom:	PD000711;	Lectin_legB; 1.
DR	PROSITE:	PS00307;	LECTIN_LEGUME_BETA; UNKNOWN_1.
KW	Lectin:	Signal,	
FT	NON_TER	1	
FT	SIGNAL	<1	
SEQ	SEQUENCE	254 AA; 28389 MW; 761b82817de7dfb9 CRC64;	
Query Match 35.5%; Score 486; DB 10; Length 254;			
Best Local Similarity 44.4%; Pred. No. 2e-29;			
Matches 111; Conservative 42; Mismatches 77; Indels 20; Gaps			
QY	3	SLSFSTKFDPNQE-DLIIFQHAT-STNNVLOVTKLDSAGNPVSAGRVLSAPRLWE	60
Db	3	SLFSFYNNFQDDERNLILQGDAKSFGIOLTKVDNDGTFPAKSTVGRLVHSTQVRLE	62
QY	61	DSA-VITSFDTIINFETISPYTSRIADGLAFFIAPPDSVISYH--GGGFLTLPNANTLNN	117
Db	63	KSTNRLTNFAQSFFVIKSPIDNG-ADGLIAFFAAFDSEIPKNSAGGTGLGF-----	113
QY	118	SSTSSENQTTHKAASSNVAVEFTYLNPDYG--DPNYIHIGIDVNSIRSKVTAKWDQMG	175
Db	114	----DPQTAQNPSANOVLAVEFTFTFAQDSNGWDPNYQHIGIDVNSIKSAATTKWERRDG	169
QY	176	KIATAHTSYNSVKRSLSVSYYAGSPATLSYDIELHTVLPEWRVGLSASTGDQKERNT	235
Db	170	QTNLNVLTYDNASKNLQVASYPDGORYQLSYRVDLDLYPEWRGVGFSAAASQQYQSHE	229
QY	236	VHWSFTSSL	245
Db	230	LQSWSETSTL	239

RESULT	313
Q9FVF8	
ID	Q9FVF8 PRELIMINARY; PRT; 258 AA.
AC	Q9FVF8;
DC	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Lectin II (Fragment).
OS	ulex europeus (furze).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eurosid I; Fabales; Fabaceae; Papilionoideae; Genisteae; Ulex.
OX	NCBI_TaxID=3902;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Loris R., De Greve H., Dao-Thi M.-H., Messens J., Imberty A., Wyns L.
RT	"Structural basis of carbohydrate recognition by lectin II from Ulex
RL	europeus, a protein with a promiscuous carbohydrate binding site.";
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF190633; AAG16779.1; .
DR	HSSP; P05048; 1SBD.
DR	InterPro; IPR000985; Lectin_legA.
DR	InterPro; IPR001220; Lectin_legB.
DR	Pfam; PF00138; lectin_legA; 1.
DR	Pfam; PF00139; lectin_legB; 1.
DR	ProDom; PD000671; Lectin_legA; 1.
DR	ProDom; PD000711; Lectin_legB; 1.
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_N_1.
FT	NON_TER
FT	1
SQ	SEQUENCE 258 AA; 27909 MW; 581F6DD8F5E049FB CRC64;

```

Query Match      33.5%; Score 459.5; DB 10; Length 258;
Best Local Similarity 42.1%; Pred. No. 2.2e-27;
Matches 107; Conservative 44; Mismatches 76; Indels 27; Gaps 10;

Qy 4 LSFSTKFDPNQEDLIFOGHAT-STNNVLQVTKLDSAGNPVSSAGRVLYSAPURLMED- 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

6 LSNFDFKVPNQKNIIFQGAASVSTTGVQTKV--SKPTTISIGRALYAAPTQIWDST 62

62 SAVLTSTDTIINFESTIPYTSRIADGLAFFIAPPDSVI--SYHGFGFLGFLFNANTLNSS 119

63 TGVASFATSFVVKADKSDGV-DGLAFFLAPANSQIPSGSSASMGCLF-----NSS 114

120 TSENQTTTAASSNVVAVEDTYLNPQYG--DPNYIIIGIDVNSIRSKVTAKWDQNGKI 177

115 DSK-----SSNQIIAVEFDYTGKAYNPWDPDFKHGIDGVNSIKSTIKTVKWDWRNGEV 167

178 ATAHISYNSYKRLSV-TSYYAGSKPATLSYDIELHTVLPEMVVGLSASTGQDKERT- 235

168 ADVVITRAPTKTSVCLSYSPDSTSNLTASVDLKAILPEWVSVGSGGVGNAAEFETH 227

236 -VHSWSTSSLTWN 248

228 DILSWYFTSNEAN 241

RESULT 14

P93246 PRELIMINARY; PRT; 256 AA.

P93246;

DT 01-MAY-1997 (TREMBlrel. 03, Created)

DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE Bark hemagglutinin (LECMABH) (Fragment).

OS Maackia amurensis.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosidis I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Maackia.

OC NCBI_Taxid=37501;

[1]

SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.

RC TISSUE=BARK;

RC MEDLINE=97390228; PubMed=9249142;

RA Van Damme E.J.M., Van Leuven F., Peumans W.J.;

RV "Isolation, characterization and molecular cloning of the bark lectins

RT from Maackia amurensis.";

RL Glycoconq. J. 14:449-456(1997).

CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS

CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE

CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY

CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS

CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS (BY SIMILARITY).

CC -!- SUBUNIT: HOMOTETRAMER OR HETEROTETRAMER OF MAHB AND MALB SUBUNITS.

CC -!- PTM: ONLY ONE OF THE PUTATIVE GLYCOSYLATION SITES IS OCCUPIED.

CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.

DR EMBL; U65008; AAB39932.1; --

DR HSPSP; P05046; LSBE.

DR InterPro; IPR000985; Lectin_legA.

DR InterPro; IPR001220; Lectin_legB.

DR Pfam; PF00138; lectin_legA;

DR Pfam; PF00139; lectin_legB; 1.

DR ProDom; PD000671; Lectin_legA; 1.

DR ProDom; PD000711; Lectin_legB; 1.

DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.

DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN.1.

KW Lectin; Calcium; Manganese; Glycoprotein; Multigene family.

FT NON_TER 1

FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 256 AA; 27840 MW; 87322A2C878459EA CRC64;

Query Match 33.5%; Score 459; DB 10; Length 256;

Best Local Similarity 43.3%; Pred. No. 2.3e-27;

Matches 116; Conservative 39; Mismatches 85; Indels 28; Gaps 12;

QY 4 LSEFSTKDPKNOEDLIFOGHAT-STNNYQVTKLDSAGNPVSSSAGRVLYSAPRLWED- 61

Db 4 LSFPIINFPVNEADLHFGVASVSTGVQLQJLS-QKNQPLEYSVGRALYSAPRVWDST 62

QY 62 SAVLTSTDTIINFESTIPYTSRI-ADGLAFFIAPPDSVISYH--GGFLGFLFNANTLNSS 118

```
Db 63 TGRVASTSTFFVQK--AARLTSGLAFLAPDSDIPSGDVSKYLGLFNNSN----- 115
QY 119 STSENQTTKAASNVAVEDTYLNDYG--DPNYIHIGIDVNSIRSKVTAKWDQWQK 176
Db 116 SESSNQ-----IVAVEFDTFENHNDPNDPNYRHIGIDVNGIDSKTKVQWDYINGG 166
QY 177 IATAHISVNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPPEWVRVGLSASTGQD--KER 233
Db 167 VAFATITLAPNKTLIASLVYPSSETSFIVAASVDLKEILPEWVRVGFSAATGAPAAAE 226
QY 234 NTVHSWSTSSLTNNAKKENKNIYTR 261
Db 227 HDVRSWSTSTFEAN-SPVDHNVHIAH 253

RESULT 15
Q8RVY4
ID Q8RVY4 PRELIMINARY; PRT; 273 AA.
AC Q8RVY4;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Phytohemagglutinin precursor.
GN LEC21.
OS Phaseolus coccineus (Scarlet runner bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3886;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Lioi L.;
RT "Lectin genes in legumes.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ438774; CAD27654.1; --
KW Signal.
FT SIGNAL 1 21 POTENTIAL
FT CHAIN 22 273 PHYTOHEMAGGLUTININ.
SQ SEQUENCE 273 AA; 29522 MW; C2FEBACFA8037F6A CRC64;

Query Match 33.2%; Score 454.5; DB 10; Length 273;
Best Local Similarity 43.4%; Pred. No. 5.6e-27;
Matches 109; Conservative 33; Mismatches 84; Indels 25; Gaps 7;

QY 1 AGSLSFSTKFDNQEDLIPOGHAT-STNNVLQVTKLDSAGNPVSSAGRVLYSAPLRWL 59
Db 22 ASSETSFSORF--NETNLIQGNASVSSGQLRLTNLKSNGEPTVGSIGRAFYSTPIQIW 79
QY 60 ED-SAVLTSFDTLINFEISTPYTSRIADGLAFTAPDPSVISYHCGFLGLFPNANTLNS 118
Db 80 DSTTGLASFATSFFTNIYAPIVKNASDGLAFALVPVGSQPKGKGGFLGLFDGSGNT--- 135
QY 119 STSENQTTKAASNVAVEDTYLNDYGDPNKITHGIDVNSIRSKVTAKWDQWQKIA 178
Db 136 -----NFTVAVEFDLYNNWN-DPKERHIGIDVNSIKSIKTTTPWDFVNGENA 182
QY 179 TATHISVNSVKRLSVTSYAGSKPA-TLSYDIELHTVLPPEWVRVGLSASTGQD---ERN 234
Db 183 KVHITYESSTKLMLASLVPSLKKSTVSDTVDLKSVLPPEWSVGVFSATTGIDKGNVETN 242
QY 235 TVHSWSTSSSL 245
Db 243 DILSWSFASKL 253
```

Search completed: February 26, 2003, 16:54:25
Job time : 33.0487 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 16:40:33 ; Search time 19.7753 Seconds
(without alignments)
1283.395 Million cell updates/sec

Title: US-09-476-485A-2

Perfect score: 1370

Sequence: 1 AGSLFSFTKFDPNQEDLIF.....LWTNVAKKENKYITRGVL 264

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	657.5	48.0	293	2 S66356	mannose/glucose-binding
2	650	47.4	290	2 A34139	concanavalin A precursor
3	640	46.7	290	1 CVJBP	concanavalin A pre
4	578	42.2	290	2 S66357	mannose/glucose-bi
5	575.5	42.0	132	2 A54864	mannose/glucose-sp
6	496	36.2	290	2 S66355	lectin-related sto
7	495	36.1	287	2 JC5444	leukoagglutinin pr
8	490	35.8	247	2 JC2268	hemagglutinin - Ma
9	478.5	34.9	286	2 S48033	lectin precursor -
10	473.5	34.6	244	2 S23099	lectin I, anti-H(O
11	463.5	33.8	285	2 S27365	lectin precursor
12	449	32.8	233	1 LNLWBA	phytohemagglutinin
13	445.5	32.5	273	2 A5701	lectin II - furze
14	441	32.2	285	2 S62690	lectin - Scotch la
15	440.5	32.0	249	2 JX0163	agglutinin II prec
16	439	32.0	250	2 S16964	probable lectin 2
17	439	32.0	285	2 S62691	mannose/glucose-sp
18	432	31.5	279	2 T09620	phytohemagglutinin
19	430.5	31.4	105	2 B4864	lectin precursor -
20	425	31.0	274	2 S1832	galactose-specific
21	423.5	30.9	249	2 JQ1981	phytohemagglutinin
22	422	30.8	275	1 LNP	lectin II - Scotch
23	420.5	30.7	276	2 S1831	lectin precursor
24	420	30.7	241	2 JX0289	phytohemagglutinin
25	420	30.7	290	2 JX0175	galactose-specific
26	419.5	30.6	281	2 S09697	lectin precursor -
27	419	30.6	241	2 JX0290	galactose-specific
28	417.5	30.5	272	2 B22826	phytohemagglutinin
29	416.5	30.4	253	2 A53416	lectin II - Lima b

30 412 30.1 233 1 FVVFBA
31 411 30.0 236 1 LNOU
32 410.5 30.0 277 2 S25296
33 405 29.6 237 2 JU0176
34 401.5 29.3 275 2 A22826
35 401 29.3 237 2 A59415
36 400 29.2 237 2 A59417
37 399.5 29.2 240 2 S36797
38 391 28.5 244 1 LNL
39 386 28.2 237 1 CVJB
40 380.5 27.8 275 2 A29572
41 375 27.4 237 2 A45587
42 364.5 26.6 275 2 A31972
43 358.5 26.2 236 2 B45587
44 353 25.8 273 2 S24044
45 345.5 25.2 243 2 JX0162

ALIGNMENTS

RESULT 1

S66356

mannose/glucose-binding lectin CLAI precursor - Cladrastis lutea

C:Species: Cladrastis lutea

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999

C:Accession: S66356; S66300

R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.

Plant Mol. Biol. 29, 579-598, 1995

A:Title: A lectin and a lectin-related protein are the two most prominent proteins in

A:Reference number: S66299; MUID:96123235; PMID:8534854

A:Accession: S66356

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-293 <VAM>

A:Cross-references: EMBL:U21958; NID:g1141756; PIDN:AAC49136.1; PID:g1141757

A:Experimental source: bark

A:Accession: S66300

A:Molecule type: protein

A:Residues: 36-46, S',48-55; D',163-179, 'X',181 <VAM>

C:Superfamily: plant lectin

C:Keywords: glycoprotein; lectin

F:1-35/Domain: signal sequence #status predicted <SIG>

F:36-293/Product: mannose/glucose-binding lectin CLAI #status experimental <MAT>

Query Match 48.0%; Score 657.5; DB 2; Length 293;

Best Local Similarity 53.2%; Pred. No. 5.3e-45;

Matches 140; Conservative 34; Mismatches 80; Indels 9; Gaps 6;

QY 3 SLFSFTKFDPNQEDLIFOGHAT-STNNVLOVTKLDSAGNPVSSSAGRVLYSAPLRWED 61

Db 38 SLSTFNFPNSEDLLIFQKDAISSNETLETRISSGQPATSVSGRALYTPVRLWDK 97

QY 62 S-AVLTPDPTTINFEISPTYSRIADGLAFFIAPPDSVISYHGGFLGFLFPNANTLNNST 120

Db 98 STGRLASFKTTFSEAITSP-TQDPGDGFAFFIAPPDTPPGYGGLLGLF---NGFNLRNS 153

QY 121 SENQTTTAAASNVAVEFDYVLPDYPNVIHIGIDVNSRSTKVTAKDWQNGKIATA 180

Db 154 SNGVAVNQAQIVAVEFDYINGQC-DPKYRHVGIDVNSITSLAYTQWQNGVKATA 212

QY 181 HISYNSVSKRLSVTSYAGSKPATLSYDIEHTVLPWVRVGLSASTGGDKERTVHWS 240

Db 213 QISYNPASOKLTAVTSYNSTPLTVSLDIDLOTVLPWVRVGFSGASTGONVERNSILAWS 272

QY 241 FTSSLTWNVAKKENKYITRGV 263

Db 273 FSSSLTTLTAKK--EDMYIARYV 293

RESULT 2

A34139

concanavalin A precursor - sword bean

A:Cross-references: EMBL:U21959; NID:g1141758; PIDN:AAC49137.1; PID:g1141759
A:Experimental source: bark
C:Superfamily: plant lectin
C:Keywords: glycoprotein; lectin
F:1-35/Domain: signal sequence #status predicted <SIG>
F:1-35/Product: mannose/glucose-binding lectin CLAI #status experimental <MAT>

Query Match 42.2%; Score 578; DB 2; Length 290;
Best Local Similarity 48.0%; Pred. No. 1.1e-38;
Matches 129; Conservative 43; Mismatches 73; Indels 24; Gaps 8;
QY 3 SLSPFTKFDNQEDLLIFQGH--TSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRWE 60
Db 38 SLSTFDNFRDQRLDLQGDKAKSSGDSQLTKTDTSGKPVRSVGRVLYSAPLRWE 97
QY 61 DSA-VLTSTFTIINFETSTPTSTRIADGLAFFIAPPDSVI--SYHGGFLGLFPNNTLNN 117
Db 98 SSTNRLASFQTFVFLSSP-TNNPGDIAFFIAPPETIPPGSSGGLGLFSPDNVNN 156
QY 118 SSTSENOTTTKAASSNVVAFEDTLYLPDYGDPNYIHIGIDVNSIRSKVTAKWDQNGKI 177
Db 157 S-----LNQIVAFEDFVNNW-DPSHRHIGIDVNTIKSSATVWQRENGSL 203
QY 178 ATAHISVNSKRLSVTSY---AGSKPATLSYDIELHTVLPWVRVGLSASTGQDKERN 234
Db 204 ATAQISVNSDTKLSVSYSPNTQANEDYTVSYDVLKTELPEWVRVGLSASTGQGVQNH 263
QY 235 TVHSWSFTSSLLWNTNVAKKENKVIITRGV 263
Db 264 NILSWTFNSNLQSSRAKK--EDIYIKRYV 290

RESULT 5
A54864
mannose/glucose-specific lectin alpha chain - Dolichos lab lab (field bean)
C:Species: Dolichos lab lab (field bean)
C:Date: 23-Jun-1995 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C:Accession: A54864; A48591
R:Gowda, L.R.; Savithri, H.S.; Rao, D.R.
J. Biol. Chem. 269, 18789-18793, 1994
A:Title: The complete primary structure of a unique mannose/glucose-specific lectin from
A:Reference number: A54864; MUID:94308133; PMID:8034631
A:Accession: A54864
A:Molecule type: protein
A:Residues: 1-132 <GOW>
C:Superfamily: plant lectin
C:Keywords: blocked amino end; glycoprotein; lectin

Query Match 42.0%; Score 575.5; DB 2; Length 132;
Best Local Similarity 91.9%; Pred. No. 6e-39;
Matches 114; Conservative 2; Mismatches 5; Indels 3; Gaps 3;

QY 130 ASSNVVAFEDT-YLPDYGDPNYIHIGIDVNSIRSKVTAKWDQNGKIATAHISVNSVS 188
Db 2 AESNVVAFEDTLYLPDYGDPNYIHIGIDVNSIRSKVTASWDQNGKIATAHISVNSVS 61
QY 189 KRLSVTSYTAG-SKPAATLSYDIELHTVLPWVRVGLSASTGQDKERNVHSWFTSSLT 247
Db 62 KRLSVTYYTPGRGKPAT-SYDIELHTVLPWVRVGLSASTGQNTERTVHSWFTSSLT 120
QY 248 NVAK 251
Db 121 NVAK 124

RESULT 6
S66355
lectin-related storage protein precursor - Cladrastis lutea (fragment)
C:Species: Cladrastis lutea
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 20-Aug-1999
C:Accession: S66355; S66299
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995

A:Title: A lectin and a lectin-related protein are the two most prominent proteins in
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66355
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-290 <VAN>
A:Cross-references: EMBL:U21940; NID:g1141754; PIDN:AAC49150.1; PID:g1141755
A:Accession: S66299
A:Molecule type: protein
A:Residues: 37-56 <VAM>
A:Experimental source: bark
C:Superfamily: This lectin-related protein has no carbohydrate binding activity.
C:Keywords: plant lectin
F:1-36/Domain: signal sequence (fragment) #status predicted <SIG>
F:37-290/Product: lectin-related storage protein #status experimental <MAT>

Query Match 36.2%; Score 496; DB 2; Length 290;
Best Local Similarity 43.2%; Pred. No. 3.9e-32;
Matches 112; Conservative 49; Mismatches 74; Indels 24; Gaps 9;
QY 3 SLSPFTKFDNQEDLLIFQGH--TSTNNVQVTKLDSAGNPVSSSAGRVLYSAPLRWE 61
Db 39 ALSFTTFKFSNQDELLQGDALVSKGELQLTRVEN-GQIPHSVGRVLYSDPVHIWDS 97
QY 62 S-AVLTSFTIINFETSTPTSTRIADGLAFFIAPPDSVISHGFLGLFPNNTLNNSS 120
Db 98 STGVSASFVTSFTFVVEAPENKNTADGIAFLAPPDQVQSLGFLGLF-NSSVYN---- 152
QY 121 SENQTTTKAASSNVVAFEDTLYLPDYGDPNYIHIGIDVNSIRSKVTAKWDQNGKIATA 180
Db 153 -----SSNQLAVEEDTFNS--WDPTARHIGIDVNSIESTRTATGWRNGEVAIV 201
QY 181 HTSYNSVSKRLSVTSYAGSKPA-TLSYDIELHTVLPWVRVGLSASTGQDK---ERNTV 236
Db 202 LITYVAPAEATLIASLTYPSSQTSYILSAVDLKSILPEWVRVGFSAATGRSAGVYETHDV 261
QY 237 HSWSTSSLLWT-NVAKKEN 254
Db 262 LSWSTSTLTETNGSGAKQN 280

RESULT 7
JC5444
leukoagglutinin precursor - Maackia amurensis
C:Species: Maackia amurensis
C:Date: 17-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 07-May-1999
C:Accession: JC5444; PC4324
R:Yamamoto, K.; Konami, Y.; Irimura, T.
J. Biochem. 121, 756-761, 1997
A:Title: Sialic acid-binding motif of Maackia amurensis lectins.
A:Reference number: JC5444; MUID:97306060; PMID:9163528
A:Accession: JC5444
A:Molecule type: mRNA
A:Residues: 1-287 <YAM>
A:Accession: PC4324
A:Molecule type: protein
A:Residues: 30-55; 74-176; 187-253 <YA2>
A:Experimental source: seed
C:Comment: This protein is a leguminous lectin. It interacts with high affinity with
te galactosyl residues.
C:Superfamily: plant lectin
C:Keywords: glycoprotein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-287/Product: leukoagglutinin #status predicted <MAT>
F:90,142,208,220/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:272/Disulfide bonds: interchain #status predicted

Query Match 36.1%; Score 495; DB 2; Length 287;
Best Local Similarity 43.1%; Pred. No. 4.6e-32;
Matches 115; Conservative 41; Mismatches 87; Indels 24; Gaps 9;

QY 4 LSPFTKFDNQEDLLIFQGHAT-STNNVQVTKLDSAGNPVSSSAGRVLYSAPLRWE- 61

[illegible]

RESULT 8
JC2268
hemagglutinin - Maackia amurensis
C:Species: Maackia amurensis
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 07-May-1999
C:Accession: JC2268
R:Konami, Y.; Ishida, C.; Yamamoto, K.; Osawa, T.; Irimura, T.
J. Biochem. 115, 767-777, 1994
A:Title: A unique amino acid sequence involved in the putative carbohydrate-binding domain of hemagglutinin (HAH)
A:Reference number: JC2268; MUID:94375425; PMID:8089095
A:Accession: JC2268
A:Molecule type: protein
A:Residues: 1-247 <KON>
A:Experimental source: seed
C:Superfamily: plant lectin
C:Keywords: calcium binding; glycoprotein; manganese
F:117,177/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:125,127,138,143/Binding site: manganese (Glu, Asp, His) #status predicted
F:127,135,138/Binding site: calcium (Asp) #status predicted

Query Match	35.8%	Score 490;	DB 2;	Length 247;
Best Local Similarity	45.7%;	Pred. No. 9.3e-32;		
Matches 116:	Conservative	34;	Mismatches 78;	Indels 26;
				Caps 10;

[illegible]

```
QY      235   TVHWSFTSSLWTN 248  
          |         |||||:|  
Db      227   DVRSWFTSTLEAN 240  
  
RESULT 9  
S48033  
lectin precursor - Robinia pseudoacacia (black locust)  
C:Species: Robinia pseudoacacia (black locust)  
C:date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Jun-2000  
C:Accession: S48033; S68376  
R:Yoshida, K.; Baba, K.; Yamamoto, N.; Tazaki, K.
```

Plant Mol. Biol. 25, 845-853, 1994
A: title: Cloning of a lectin cDNA and seasonal changes in levels of the lectin and its mRNA in *Lotus corniculatus*
A: reference number: S48033; MUID: 94355657; PMID: 7915553
A: accession: S48033
A: molecule type: mRNA
A: residues: 1-286 <YOS>
A: cross-references: EMBL:D17757; NID:g538528; PID:g538529
A: experimental source: inner bark
R: tazaki, K.; Yoshida, K.; Shinohara, K.; Koshiba, T.; Yamamoto, N.

Query Match	34.9%	Score 478.5;	DB 2;	Length 286;
Best Local Similarity	44.7%;	Pred. No. 9.4e-31;		
Matches 117;	Conservative	40;	Mismatches 80;	
			Indels 25;	Gaps 10;

QY	2	GSLSFSTKFDPNQEDLIFOGHAISTN-NVLQVTKLDISAGNPVSSAGSRVLYSAPRLWE	60
		: : : : : :	
Dd	33	GSLSFSPFKPHSQDPLITQSADALVTSGVKQLTUTTYND-GRPVYIGRVLAAFPQIWD	91
		: : : : : :	
QY	61	D-SAVLTSDFTIINFELSTPYTSRIADGLGAFFTAPPDSVISYH-GGFLGLEFNANTLNNS	118
		: : : : : :	
Dd	92	STTGCVNASFVTSFIIKAPNECKTAGLVFFLPVGSVQLKGGLLGLF-----	142
		: : : : : :	
QY	119	STSENOTTTKAASNVAVEFDLYLNPDYGDPNYIHIGIDVNSIRSKVTAKWDMQNGKTA	178
		: : : : : : : : : : : : : : : : : : : :	
Dd	143	--KDESYNK--SNQIVAVEFDTFERNVAW-DPNGIHMGIHDVNSIQSVRTVRMDWANGEYA	196
		: : : : : :	
QY	179	TAMHSNVSKRLSVTSYYAG-SKPATLSVDIELHVLPPEWRVGLSASTGDK---ERN	234
		: : : : : :	
Dd	197	NVFIYSFASTRKSUTASLVYPSELSEFILSAIVDELKKVLPEWRVVGFTATTGLSEDIYQVTN	256
		: : : : : :	
QY	235	TVHSWSFTSSL--WTNVAKEN	254
		: : : : : :	
Dd	257	DVLSWSFESNLPGGNSVASVKN	278
		: : : : : :	

RESULT 10
S23099 lectin I, anti-H(O) - Cytisus sessilifolius
C:Species: Cytisus sessilifolius
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-May-1998
C:Accession: S23099; SL3438
R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.
FEBS Lett. 304, 129-135, 1992
A:Title: Correlation between carbohydrate-binding specificity and amino acid sequence
A:Reference number: S23099; MUID:92316214; PMID:1618311
A:Accession: S23099
A:Molecule type: protein
A:Residues: 1-244 <KON>
A:Experimental source: seed
R:Konami, Y.; Yamamoto, K.; Osawa, T.
Biol. Chem. Hoppe-Seyler 372, 103-111, 1991
A:Title: Purification and characterization of two types of Cytisus sessilifolius anti-
A:Reference number: S13438; MUID:91315748; PMID:1859626
A:Accession: S13438

A: Molecule type: protein
 A: Residues: 1-3, 'K', 5-23, 'A', 25-29, 'K', 31-35 <BIO>
 A: Experimental source: seed
 C: Superfamily: plant lectin
 C: Keywords: calcium binding; glycoprotein; homotetramer; lectin; manganese; seed
 F: 115-129/Domain: glycopeptide #status predicted <GUP>
 F: 117,117/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F: 127,129,140,145/Binding site: manganese (Glu, Asp, Asp, His) #status predicted

F;129,140/Binding site: calcium (Asp) #status predicted

```

Query Match          34.6%;      Score 473.5;   DB 2:   Length 244;
Best Local Similarity 42.4%;      Pred. No. 1.9e-30;
Matches 108; Conservative 43; Mismatches 81; Indels 23; Gaps

QY      4  LSFSEFTKFDPNQEDLFIQGHAT-STNNVLQVTKLDAGNPNVSSAGKRVLYSAPFLRLWED- 61
      Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      5  LSFNFKEFVFNQNNLFIQGEASVSTTGVQLQVTKV---SKPATHSIGRALYAAPVHLDWST 61

QY      62  SAVLTGSDFTIINFEI-STPYTSRIADGLAFFIAPPDSVSYHGGFTLGLFPNANTLNSST 120
      Db      : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      62  TGRVASGETSFVFWKDEPEKSGVGLTFFFLAPANSQIP-SGSSAGLFLGFLNSSDNKS- 119

QY      121  SENQTTTKAASNVAVEFDITVLPNDYG--DPNYTHIGIDVNSIRSKVTAKWDWQNGKIA 178
      Db      | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      120  -----SNGIIAVEFDYFGKTYNPWDPDFKHIQVDVNSIKSIKTVKWDWRNGEVA 169

QY      179  TAHISYNSYSKRLSVSYVYAGSKPATL-SYDTELHTVLPPEWVRVGLSASTGQDKERNT-- 235
      Db      | : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      170  NVVITYRAPTKSLTSLVSPSDQTSNIVTASVDLKKAILPEWVSVGFGVGNAAEFETHD 229

QY      236  VHSWSTSSLWNTVA 250
      Db      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
      230  VLSWYFTSLEANPA 244

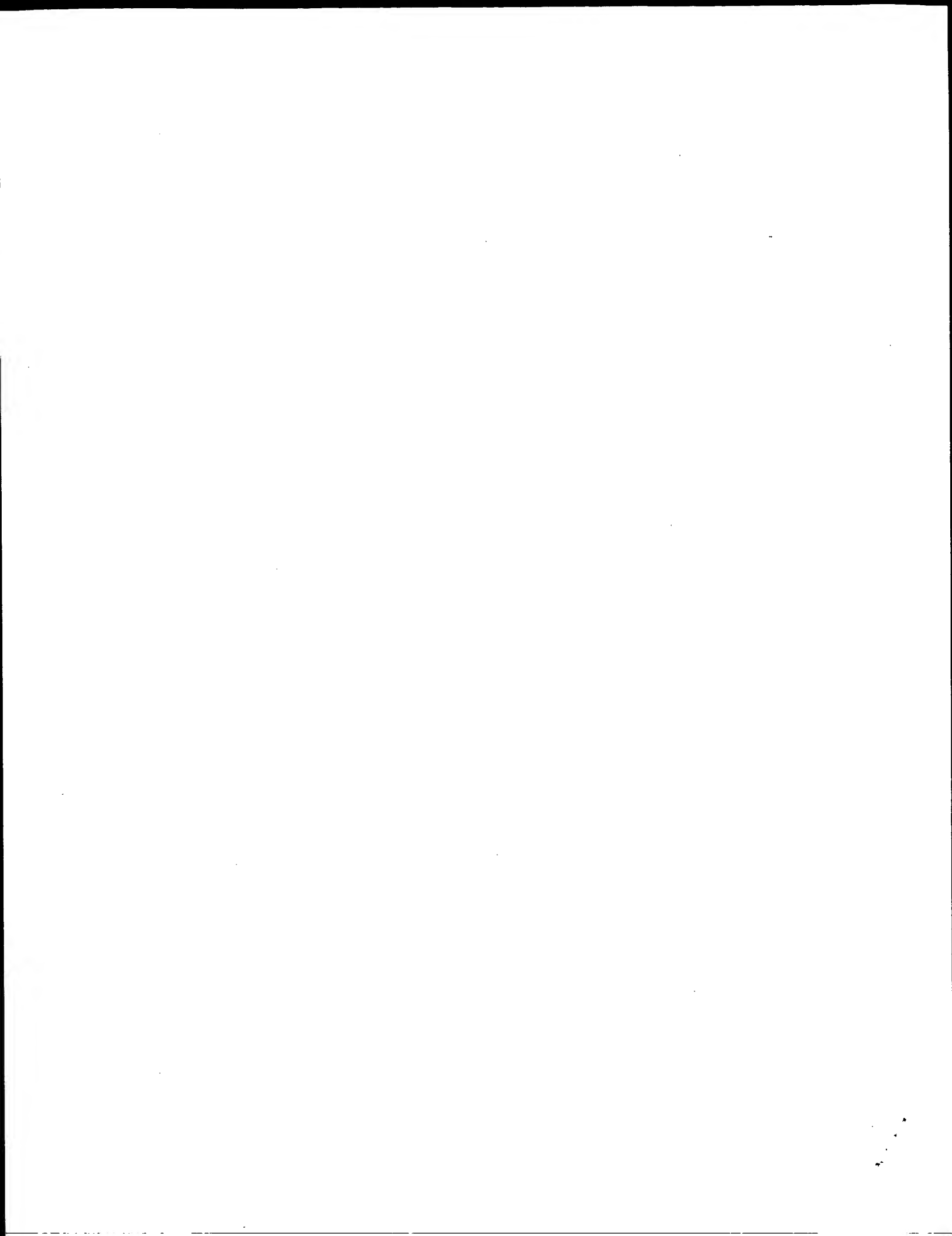
RESULT 11
S27365
lectin precursor - soybean
C:Species: Glycine max (soybean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C:Accession: S27365
R:Vodkin, L.O.; Rhodes P.R.; Goldberg, R.B.

```

[illegible]

OY 233 RNFVHWSFTSSLWTN 248
: | | | | | | | | | |
Db 230 DHDVLSHWFTTSNLEAN 245

Search completed: February 26, 2003, 16:51:31
Job time : 20.7753 secs




```

||||| 166 TATAHLSYNSASKRLSSVSSYPNVVVSFDVLENNVGPDPVVRVGSATTTGTYTQNNI 225
QY 237 HSWSTSSL 245
DB 226 LAWSFRSSL 234

RESULT 5
CONA_CANEN STANDARD; PRT; 290 AA.
AC P02866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Concanavalin A precursor (Con A).
OS Canavalia ensiformis (Jack bean) (Horse bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Canavalia.
OX NCBI_TaxID=3823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85086270; PubMed=3965973;
RA Carrington D.M., Auffret A., Hanke D.E.;
RT "Polypeptide ligation occurs during post-translational modification
of concanavalin A."
RN [2]
RP NATURE 313:64-67(1985).
RN [2]
RP SEQUENCE OF 30-148.
RX MEDLINE=75095622; PubMed=1112813;
RA Wang J.L., Cunningham B.A., Waxdal M.J., Edelman G.M.;
RT "The covalent and three-dimensional structural of concanavalin A. I.
Amino acid sequence of cyanogen bromide fragments F1 and F2."
RN [3]
RP J. Biol. Chem. 250:1490-1502(1975).
RN [3]
RP SEQUENCE OF 164-281.
RX MEDLINE=75095623; PubMed=1112814;
RA Cunningham B.A., Wang J.L., Waxdal M.J., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. II.
Amino acid sequence of cyanogen bromide fragment F3."
RN [4]
RP J. Biol. Chem. 250:1503-1512(1975).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RX MEDLINE=73053316; PubMed=4638345;
RA Hardman K.D., Ainsworth C.F.;
RT "Structure of concanavalin A at 2.4-A resolution."
RN [5]
RP Biochemistry 11:4910-4919(1972).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=75095624; PubMed=1112815;
RA Becker J.W., Reeke G.N. Jr., Wang J.L., Cunningham B.A., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. III.
Structure of the monomer and its interactions with metals and
saccharides."
RN [6]
RP J. Biol. Chem. 250:1513-1524(1975).
RN [6]
RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=75095625; PubMed=1112816;
RA Reeke G.N. Jr., Becker J.W., Edelman G.M.;
RT "The covalent and three-dimensional structure of concanavalin A. IV.
Atomic coordinates, hydrogen bonding, and quaternary structure."
RN [7]
RP J. Biol. Chem. 250:1525-1547(1975).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
RA Parkin S., Rupp B., Hope H.;
RT "Atomic resolution structure of concanavalin A at 120 K."
RN [8]
RP Acta Crystallogr. D 52:1161-1168(1996).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
RX MEDLINE=20347885; PubMed=10748006;
RA Bouckaert J., Dewallef Y., Poortmans F., Wyns L., Loris R.;
RT "The structural features of concanavalin A governing non-proline
peptide isomerization."
RN [9]
RP J. Biol. Chem. 275:19778-19787(2000).
RN [9]
RP FUNCTION: D-mannose specific lectin.
CC -!- SUBUNIT: Homotetramer.
CC -!- PM: THE MATURE CHAIN CONSISTS OF RESIDUES 164-281 FOLLOWED BY
POST-TRANSLATIONAL MODIFICATION AFTER REMOVAL OF THE SIGNAL
SEQUENCE; CLEAVAGE AFTER ASN AT POSITIONS 148, 163, AND 281 IS
FOLLOWED BY TRANSPOSITION AND LIGATION (BY FORMATION OF A NEW
PEPTIDE BOND) OF RESIDUES 164-281 AND 30-148.
CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
SACCHARIDE-BINDING AND CELL-AGGLUTININATING ACTIVITIES.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -!- DATABASE: NAME=worthington enzyme manual;
WWW="http://www.worthington-biochem.com/manual/C/CONA.html".
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X01632; CAA25787.1; -
CC PIR; A03357; CVJBP.
CC PIR; A03358; CVJBP.
CC PDB; 1CNI; 30-SEP-83.
CC PDB; 2CNA; 31-JUL-94.
CC PDB; 3CNA; 15-APR-91.
CC PDB; 5CNA; 15-OCT-94.
CC PDB; 1CON; 31-JAN-94.
CC PDB; 2CTV; 31-OCT-93.
CC PDB; 1SCR; 15-OCT-94.
CC PDB; 1SGR; 15-OCT-94.
CC PDB; 1APN; 03-APR-96.
CC PDB; 1CES; 15-FEB-97.
CC PDB; 1CJP; 15-OCT-97.
CC PDB; 1CVN; 14-OCT-96.
CC PDB; 1ENQ; 17-AUG-96.
CC PDB; 1ENR; 17-AUG-96.
CC PDB; 2ENR; 16-FEB-99.
CC PDB; 1ENS; 17-AUG-96.
CC PDB; 1GIC; 20-AUG-97.
CC PDB; 1JBC; 12-FEB-97.
CC PDB; 1NLS; 26-NOV-97.
CC PDB; 1ONA; 17-SEP-97.
CC PDB; 1TEI; 24-JUN-98.
CC PDB; 1VAL; 11-JAN-97.
CC PDB; 1VAM; 11-JAN-97.
CC PDB; 1VLN; 01-APR-97.
CC PDB; 1BXH; 07-OCT-98.
CC PDB; 2CAU; 25-NOV-98.
CC PDB; 2CAV; 25-NOV-98.
CC PDB; 1DQ0; 19-JAN-00.
CC PDB; 1DQ1; 19-JAN-00.
CC PDB; 1DQ2; 19-JAN-00.
CC PDB; 1DQ4; 19-JAN-00.
CC PDB; 1DQ5; 19-JAN-00.
CC PDB; 1DQ6; 19-JAN-00.
CC InterPro; IPR000985; Lectin_lega.
CC InterPro; IPR001220; Lectin_legB.
CC Pfam; PF00138; lectin_legA; 1.
CC Pfam; PF00139; lectin_legB; 1.
CC ProDom; PD000671; Lectin_legA; 1.
CC ProDom; PD000711; Lectin_legB; 1.
CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC Lectin; Calcium; Manganese; Glycoprotein; Signal; 3D-structure.
CC SIGNAL
CC CHAIN
FT SIGNAL 1 29
FT CHAIN 30 148
FT PROPEP 149 163

```

FT CHAIN 164
 FT PROPEP 282
 FT SITE 148
 FT SITE 149
 FT SITE 163
 FT SITE 164
 FT METAL 281
 FT METAL 171
 FT METAL 173
 FT METAL 175
 FT METAL 177
 FT METAL 182
 FT METAL 187
 FT METAL 187
 FT CARBOHYD 152
 FT CONFLICT 33
 FT CONFLICT 35
 FT CONFLICT 35
 FT CONFLICT 62
 FT CONFLICT 66
 FT CONFLICT 98
 FT CONFLICT 100
 FT CONFLICT 107
 FT STRAND 167
 FT HELIX 178
 FT TURN 181
 FT STRAND 187
 FT STRAND 199
 FT TURN 207
 FT STRAND 210
 FT TURN 219
 FT STRAND 223
 FT TURN 231
 FT STRAND 236
 FT HELIX 244
 FT TURN 247
 FT STRAND 251
 FT STRAND 266
 FT TURN 283
 FT STRAND 286
 FT STRAND 289
 SQ SEQUENCE 290 AA; 31521 MW; 66CD1C62201720DD CRC64;

Query Match 46.7%; Score 640; DB 1; Length 290;
 Best Local Similarity 53.8%; Pred. No. 1.5e-43;
 Matches 134; Conservative 34; Mismatches 75; Indels 6; Gaps 4;

QY 3 SLFSFTKEDPNQEDLIFQGHATS--TNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWED 61
 Db ALHFMNQSKDKDLILQGDATFTEGRLRLTRVSNQSGSSVGRALFYAPVHIWES 95
 QY 62 SAVLTSTFTIINFEISTPTTSRIADGLAFIAPPDSVI--SYHGFGFLGPPNANTLANS 119
 Db SAVVASFEATFTFLIKSP-DSHPADGIAFIISIDSSIPSGSTGRLLGLFPDANVIRNST 154
 QY 120 TSENQTTTAAASNVAVEFDYVLPDYPNYIHIGIDVNSIRSKVTAKWQNGKIAT 179
 Db TIDFNAAYNA--DTIIVAVELDTYPTNDIGDPSYPHIGIDIKSVRSKKTAKWQNGKAVGT 212
 QY 180 AHISYNSVSKRLSVTSYXAGSKRPATLSYDIELHTVLPWVRVGLSASTGDKERTVHWS 239
 Db AHIIYNSVDKRLSASVSYNPADSATSVYDVLNDVLPWVRVGLSASTGLYKTNLTLSW 272
 QY 240 SFTSSLWTN 248
 Db 273 SFTSKLSN 281

RESULT 6
 LEC2_CLALU
 ID LEC2_CLALU STANDARD; PRT; 290 AA.
 AC Q39529;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Agglutinin II precursor (ClalII) (LeccIIAII).
 OS Cladrastis lutea (yellow wood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC CONCAVALIN (FIRST PART).
 OC CLEAVAGE.
 OC CLEAVAGE.
 OC CLEAVAGE.
 OC MANGANESE.
 OC MANGANESE AND CALCIUM.
 OC CALCIUM.
 OC CALCIUM.
 OC MANGANESE AND CALCIUM.
 OC MANGANESE.
 OC N-LINKED (GLCNAC. . .).
 OC E -> Q (IN REF. 2).
 OC N -> D (IN REF. 2).
 OC E -> D (IN REF. 3).
 OC R -> E (IN REF. 3).
 OC V -> T (IN REF. 3).
 OC AS -> SA (IN REF. 3).
 OC T -> A (IN REF. 3).
 OC STRAND 167
 OC TURN 173
 OC HELIX 178
 OC TURN 181
 OC STRAND 187
 OC STRAND 192
 OC STRAND 199
 OC TURN 202
 OC STRAND 207
 OC TURN 210
 OC STRAND 218
 OC TURN 219
 OC STRAND 222
 OC TURN 223
 OC STRAND 229
 OC TURN 231
 OC STRAND 232
 OC TURN 236
 OC STRAND 241
 OC TURN 244
 OC STRAND 247
 OC TURN 251
 OC STRAND 259
 OC TURN 266
 OC STRAND 280
 OC TURN 283
 OC STRAND 286
 OC STRAND 289
 SQ SEQUENCE 290 AA; 31521 MW; 66CD1C62201720DD CRC64;

Query Match 46.7%; Score 640; DB 1; Length 290;
 Best Local Similarity 53.8%; Pred. No. 1.5e-43;
 Matches 134; Conservative 34; Mismatches 75; Indels 6; Gaps 4;

QY 3 SLFSFTKEDPNQEDLIFQGHATS--TNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWED 61
 Db ALHFMNQSKDKDLILQGDATFTEGRLRLTRVSNQSGSSVGRALFYAPVHIWES 95
 QY 62 SAVLTSTFTIINFEISTPTTSRIADGLAFIAPPDSVI--SYHGFGFLGPPNANTLANS 119
 Db SAVVASFEATFTFLIKSP-DSHPADGIAFIISIDSSIPSGSTGRLLGLFPDANVIRNST 154
 QY 120 TSENQTTTAAASNVAVEFDYVLPDYPNYIHIGIDVNSIRSKVTAKWQNGKIAT 179
 Db TIDFNAAYNA--DTIIVAVELDTYPTNDIGDPSYPHIGIDIKSVRSKKTAKWQNGKAVGT 212
 QY 180 AHISYNSVSKRLSVTSYXAGSKRPATLSYDIELHTVLPWVRVGLSASTGDKERTVHWS 239
 Db AHIIYNSVDKRLSASVSYNPADSATSVYDVLNDVLPWVRVGLSASTGLYKTNLTLSW 272
 QY 240 SFTSSLWTN 248
 Db 273 SFTSKLSN 281

RESULT 6
 LEC2_CLALU
 ID LEC2_CLALU STANDARD; PRT; 290 AA.
 AC Q39529;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Agglutinin II precursor (ClalII) (LeccIIAII).
 OS Cladrastis lutea (yellow wood).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.
 OC NCBI_TaxID=38412;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 36-54.
 RC TISSUE=Bark;
 RX MEDLINE=96123235; PubMed=8534854;
 RA van Damme E.J.M., Barre A., Rouge P., van Leuven F.,
 RA Peumans W.J.;
 RT "A lectin and a lectin-related protein are the two most prominent
 RT proteins in the bark of yellow wood (Cladrastis lutea).";
 RL Plant Mol. Biol. 29:579-598(1995)
 CC -!- FUNCTION: MANNANOSE-BINDING BARK LECTIN.
 CC -!- FUNCTION: BARK LECTINS ARE STORAGE PROTEIN THAT PROBABLY MAINTAINS
 CC STOCKS OF NITROGEN DURING DORMANT PERIOD. SELF-AGGREGATABLE
 CC MOLECULES THAT CAN BIND THEIR OWN CARBOHYDRATE SIDE CHAINS. THEY
 CC COULD ALSO PLAY A ROLE IN THE PLANT'S DEFENCE AGAINST PHYTOPHAGOUS
 CC INVERTEBRATES OR HERBIVOROUS HIGHER ANIMALS.
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- MISCELLANEOUS: BINDS ONE MANGANESE (OR OTHER TRANSITION METAL) ION
 CC AND ONE CALCIUM ION. THE METAL IONS ARE ESSENTIAL FOR THE
 CC SACCHARIDE-BINDING AND CELL-AGGLUTINATING ACTIVITIES (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U21959; AAC49137.1; -;
 CC HSSP; P02866; 1DQ2.
 CC InterPro; IPR000985; Lectin_legA.
 CC InterPro; IPR001220; Lectin_legB.
 CC Pfam; PF00138; lectin_legA; 1.
 CC Pfam; PF00139; lectin_legB; 1.
 CC ProDom; PD000671; Lectin_legA; 1.
 CC ProDom; PD000711; Lectin_legB; 1.
 CC PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
 CC PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
 CC Lectin; Calcium; Manganese; Glycoprotein; Signal; Mannose-binding.
 CC SIGNAL 1 35
 CC CHAIN 36 290 AGGLUTININ II.
 CC METAL 165 165 MANGANESE (BY SIMILARITY).
 CC METAL 167 167 MANGANESE AND CALCIUM (BY SIMILARITY).
 CC METAL 171 171 CALCIUM (BY SIMILARITY).
 CC METAL 175 175 MANGANESE AND CALCIUM (BY SIMILARITY).
 CC METAL 180 180 MANGANESE (BY SIMILARITY).
 CC CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 200 200 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 47 51 RPDQR -> SPNEA (IN REF. 1; AA SEQUENCE).
 SQ SEQUENCE 290 AA; 32003 MW; 85060A0EEA246A0B CRC64;

Query Match 42.4%; Score 581; DB 1; Length 290;
 Best Local Similarity 48.3%; Pred. No. 6.9e-39;
 Matches 130; Conservative 42; Mismatches 73; Indels 24; Gaps 8;

QY 3 SLFSFTKEDPNQEDLIFQGHATS--TNNVLOVTKLDSAGNPVSSAGRVLYSAPRLWED 60
 Db 38 SLSTFDNFRPDQRLILQGDAKISSGDSLQTLTKDTSGKPVGRVSGRALLYTPHLWD 97
 QY 61 DSA-VLTSFDTIINFEISTPTTSRIADGLAFIAPPDSVI--SYHGFGFLGPPNANTLNN 117
 Db 98 SSTNRLASFQTTFTFLVSSP-TNNPCDGIATFIAPPETTIPFGSSGGLGLFSPDNALNN 156
 QY 118 SSTSENQTTTAAASNVAVEFDYVLPDYPNYIHIGIDVNSIRSKVTAKWQNGKI 177
 Db 157 S-----LNQIVAVEFDTEFNNW-DPSHRHIGIDVTNTIKSSATVWRQENGSL 203
 QY 178 ATAHISYNSVSKRLSVTSY---AGSKPATLSYDIELHTVLPWVRVGLSASTGDKERN 234

72

72

72

72

72

LECTIN-RELATED PROTEIN.


```

Db 92 STTGNAVSEVTSFSLIKAPNEGKTADGLVFLAPVGTQPLKGGGLGLF----- 142
QY 119 STSENQTTTAASSNVAVVEFTYLPDPYDNPYTHIGIDVNSTRSKYTAKWDWONGKTA 178
Db 143 ---KDESNK--SNOIVAVEEDTFRNVAW--DFNGIHMGIDVNSIQSVTRVRWDWANGGEVA 196
QY 179 TAHLSYNSVKRLSVTSYIAG--SKPATLSYDIELHTVLPEWVRVGLSASTGDK---ERN 234
Db 197 NVFTSYEASTKSLTASLVYPSLEKSFILSAIVDLKKVLPEWVRVGTFTAITGLSEDIYVQTN 256
QY 235 TVHSWSPTSSL--WTNVAKKEN 254
Db 257 DVLWSFESNLPGGNSVASVKN 278

RESULT 12
LEC_SOYBN STANDARD; PRT; 285 AA.
AC P05046;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Lectin precursor (Agglutinin) (SBA).
GN LEL.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=84026469; PubMed=6313203;
RA Vodka L.O., Rhodes P.R., Goldberg R.B.;
RT "CA lectin gene insertion has the structural features of a
RT transposable element.";
RN [2]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=95226380; PubMed=7711015;
RA Dessen A., Gupta D., Sabesan S., Brewer C.F., Sacchettini J.C.;
RT "X-ray crystal structure of the soybean agglutinin cross-linked with
RT a biotennary analog of the blood group I carbohydrate antigen.";
RL Biochemistry 34:4933-4942(1995).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.52 ANGSTROMS).
RX MEDLINE=98060759; PubMed=9398234;
RA Olsen L.R., Dessen A., Gupta D., Sabesan S., Sacchettini J.C.,
RA Brewer C.F.;
RT "X-ray crystallographic studies of unique cross-linked lattices
RT between four isomeric biantennary oligosaccharides and soybean
RT agglutinin.";
RL Biochemistry 36:15073-15080(1997).
CC !- FUNCTION: BINDS GALNAc AND GALACTOSE.
CC !- SUBUNIT: HOMOTETRAMER.
CC !- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; K00821; AAA33983.1; --
DR PIR; S27365; S27365.
DR PDB; 2SBA; 09-DEC-98.
DR PDB; 1SBD; 22-APR-98.
DR PDB; 1SBE; 22-APR-98.
DR PDB; 1SBF; 22-APR-98.
DR GlycoSuiteDB; P05046; --
DR InterPro; IPR000985; Lectin_legA.

```

```

DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUM_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUM_ALPHA; 1.
KW Lectin; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 32
FT CHAIN 33 285 LECTIN.
FT CARBOHYD 107 107 N-LINKED (GLCNAC...).
SQ SEQUENCE 285 AA; 30928 MW; B3704533C9315C52 CRC64;

Query Match 33.8%; Score 463.5; DB 1; Length 285;
Best Local Similarity 42.8%; Pred. No. 1.2e-29;
Matches 107; Conservative 39; Mismatches 81; Indels 23; Gaps 7;

QY 1 AGSLSFSTKFDPNQEDLIFQGHATSTNN-VLQVTKLDSAGNPVSSSAGRVLYSAPLRMW 59
Db 33 AETVFSWNKFKVPQNMILQGDVITSSGKLQNKVDENGTPKPSLGRALYSTPIHIW 92
QY 60 E-DSAVLTISFDITINFEISTPYTSRIADGLAFFIAPPDSVISYHGGFLGLFPNPANTLNS 118
Db 93 DKETGSVASEFAASFNETFYAPDTKRLADGLAFFLAPIDTKPQTHAGYGLF----- 143
QY 119 STSENQTTTAASSNVAVVEFTYLPDPYDNPYTHIGIDVNSTRSKYTAKWDWONGKIA 178
Db 144 ---KDESNK--SNOIVAVEEDTFRNVAW--DFNGIHMGIDVNSIQSVTRVRWDWANGKVA 194
QY 179 TAHLSYNSVKRLSVTSYIAG--SKPATLSYDIELHTVLPEWVRVGLSASTGDK---KERNT 235
Db 195 KVLITYDASTLSLVASLVYPSQTSNLSLDVVDLTKLSLPEWVRIGFSAATGLDIPGESH 254
QY 236 VHSWSPTSSL 245
Db 255 VLSWSFASNL 264

RESULT 13
LEC4_GRISI STANDARD; PRT; 243 AA.
AC P24146;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Lectin IV (GS4).
OS Griffonia simplicifolia.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Caesalpinoideae; Cerciloideae;
OC Griffonia.
OX NCBI_TaxID=3850;
RN [1]
RN SEQUENCE.
RX MEDLINE=93240544; PubMed=8478943;
RA Smillie L.B.;
RL Unpublished results, cited by:
RL Delbaere L.T.J., Vondonselaar M., Prasad L., Quail J.W.,
RL Wilson K.S., Dauter Z.;
RN J. Mol. Biol. 230:950-965(1993).
RN [2]
RN SEQUENCE OF 1-62.
RC TISSUE-Seed;
RX MEDLINE=91097444; PubMed=2268264;
RA Nikrad P.V., Pearlstone J.R., Carpenter M.R., Lemieux R.U.,
RA Smillie L.B.;
RT "Molecular-mass heterogeneity of Griffonia simplicifolia lectin IV
RT subunits. Differences in the oligosaccharide moieties in the
RT N-terminal region.";
RN Biochem. J. 272:343-350(1990).
RN [3]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=93240544; PubMed=8478943;

```



```

DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin: Manganese; Calcium; Signal; Multigene family; Glycoprotein.
FT SIGNAL 1 26
FT CHAIN 27 280
FT METAL 148 148
FT METAL 150 150
FT METAL 152 152
FT METAL 154 154
FT METAL 158 158
FT METAL 170 170
FT METAL 176 176
FT CARBOHYD 163 163
FT CARBOHYD 272 272
SQ SEQUENCE 280 AA; 30473 MW; DB68690AD8015E81 CRC64;

Query Match 32.78; Score 448; DB 1; Length 280;
Best Local Similarity 39.98; Pred. No. 2e-28;
Matches 109; Conservative 42; Mismatches 84; Indels 38; Gaps 9;

QY 1 AGSLSFSTKFPDQEDLIFOGHA--TSTNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRMW 59
| : ||| ||| | : ||| ||| | : ||| ||| | : ||| ||| | : ||| ||| | :
27 AETTSFSTKFPVPOKNLIFOGDAKTASTGKLESK-----AVKNSIGRALYSAPIHW 80

QY 60 ED-SAVLTSEDITINFEISTPYTSRIADGLAFFIAP-----PDSVSYHGGFLGLFPNANT 114
: : : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
81 DSKTGSVANFQTTFTTITAPNTYNAVADGLAFFIAPIDTKPKSI--HHGGYLGVF----- 133

QY 115 LNNSTSENQTTTKAASNVAVEFDTVLPNDYDGNP-----YIHIGIDVNSIRSKVT 167
: : : : : | : ||| ||| | : ||| ||| ||| ||| ||| ||| ||| ||| |||
134 -----DSKTYKKSITQTVAVEIDTFYNAQW--DPNFGNISTSGRHHGIDVNSIKSIST 183

QY 168 AKWDQWKGKATAHISYNSVKRSLSVTSYYAGSKPATLSYDIELHTVLPWVRVGLSAST 227
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
184 VPWSLENNKKNANVAIGENGADNVLSDVDEYPLRIHYTLSHVPLKDVVPWVRIGFSST 243

QY 228 GQDKERTVHSWTSSTSLTWTNVAKENKIYT 260
| : : : ||| ||| | : | : | : | : | : | : | : | : | : | : | :
244 GAEYSAHDLSWSPDSKL--NLGFENNINANVS 274

RESULT 15
PHAM_PHAVU
ID PHAM_PHAVU STANDARD; PRT; 273 AA.
AC FI5231;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Leucoagglutinating phytohemagglutinin precursor (PHA-L).
GN PDLEC2.
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. PINTO UI111; TISSUE=Leaf;
RA Voelker T.A., Staswick P., Chrispeels M.J.;
RT "Molecular analysis of two phytohemagglutinin genes and their
RT expression in Phaseolus vulgaris cv. Pinto, a lectin-deficient
RT cultivar of the bean.";
RL EMBO J. 5:3075-3082(1986).
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X04659; CAA28362.1; -.
DR PIR: A25701; A25701.
DR HSSP; P05087; IFAT.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Signal; Glycoprotein; Multigene family.
FT SIGNAL 1 21
FT CHAIN 22 273
FT CARBOHYD 33 33
FT CARBOHYD 92 92
FT CARBOHYD 92 92
SQ SEQUENCE 273 AA; 29421 MW; A148359D49538EC3 CRC64;

Query Match 32.5%; Score 445.5; DB 1; Length 273;
Best Local Similarity 42.6%; Pred. No. 3.1e-28;
Matches 107; Conservative 36; Mismatches 83; Indels 25; Gaps 7;

QY 1 AGSLSFSTKFPDQEDLIFOGHAT--STNNVLQVTKLDSAGNPVSSSAGRVLYSAPLRMW 59
| ||| | : | : ||| | : | : ||| | : | : ||| | : | : ||| | : | : ||| | :
22 ASQTFESFDRF--NETNLIQGDASVSSGQLRLTNVNSGEPVGSIGRAFYSAPIQIW 79

QY 60 E-DSAVLTSEDITINFEISTPYTSRIADGLAFFIAPDPSVSYHGGFLGLFPNANTLNS 118
: : : ||| | : | : ||| | : | : ||| | : | : ||| | : | : ||| | : | : ||| | :
80 DYTGNVASPDFTNFTNLVNNAGPADGLAFALVPVGSQPKDKGFLGLFDGSN----- 134

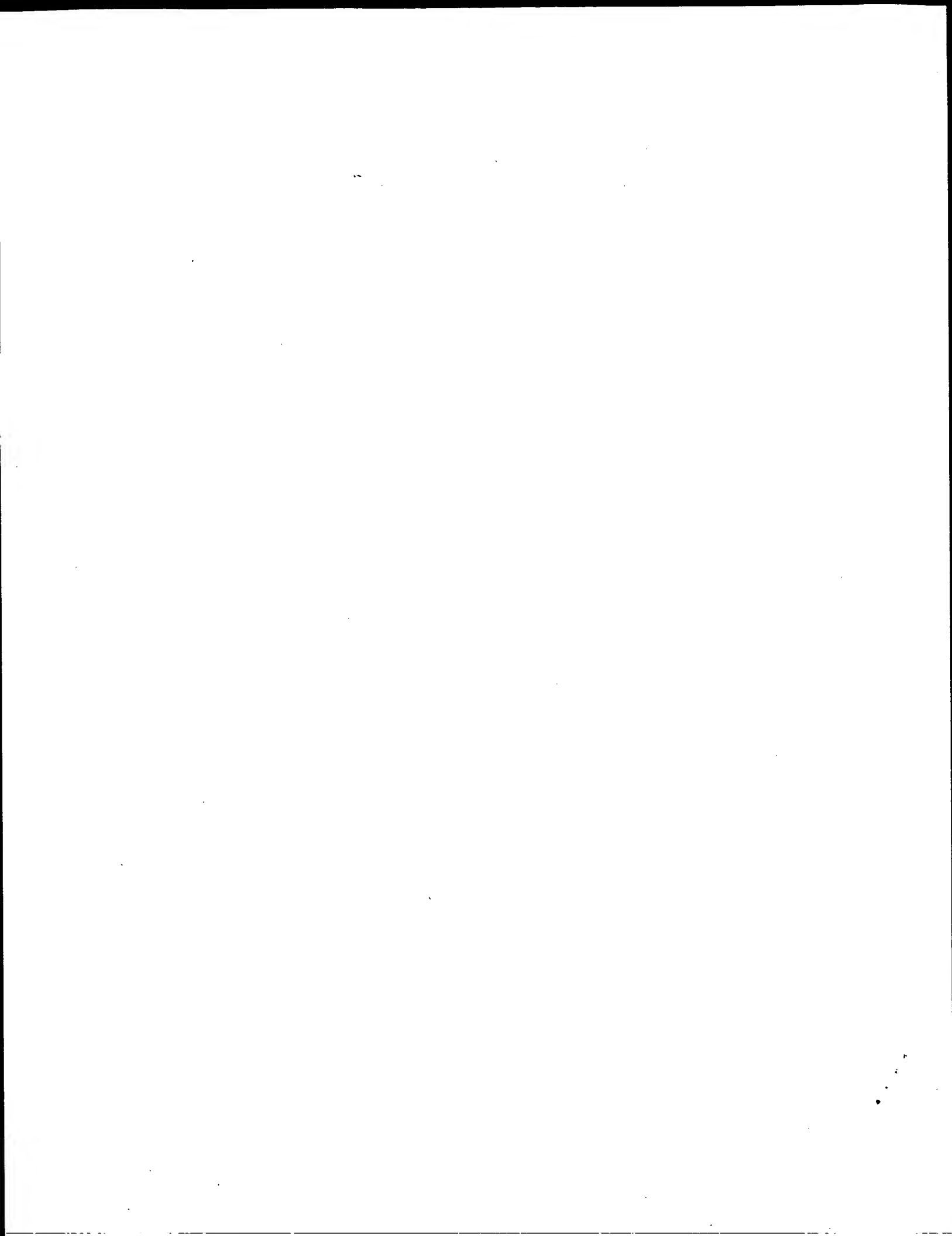
QY 119 STSENQTTTKAASNVAVEFDTVLPNDYDGNPNYIHIGIDVNSIRSKVTAKWDQWNGKIA 178
: : : ||| ||| | : | : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
135 -----SNFHTVAVEFDLYLNKDW--DPRRHGIDVNSIKSIKTTWDFVNGENA 182

QY 179 TAHISTNSVSKRSLSVTSYYAGSKPA--TSLYDIELHTVLPWVRVGLSASTQDCK---ERN 234
||| | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
183 EVHITYESSTKLIVASLVYPSLKTSTFTVSDTVLKSVLPEWVSFGFSATTGITKGNVETN 242

QY 235 TVHSWSFTSSL 245
: : ||| ||| |
DB 243 DILSWSFASKL 253

```

Search completed: February 26, 2003, 16:52:55
Job time : 9.58052 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:36:06 ; Search time 16 seconds
(without alignments)
48.067 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_73.*

2: piri.*

3: piri3.*

4: piri4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	83.3	115	2 A71809	hypothetical prote
2	30	83.3	115	2 D64710	hypothetical prote
3	30	83.3	248	2 T26523	hypothetical prote
4	30	83.3	345	2 A97252	sporulation protei
5	30	83.3	544	2 F81059	YibX/YhjW/YijP/Yjd
6	30	83.3	822	2 T41941	glycoprotein B - h
7	30	83.3	960	2 G84852	probable receptor-
8	29	80.6	311	2 G87273	oxidoreductase iro
9	29	80.6	374	2 A46352	ORF1 protein - chl
10	29	80.6	407	2 T04199	hypothetical prote
11	29	80.6	501	2 T43047	retrovirus-related
12	29	80.6	524	2 H84510	probable receptor-
13	29	80.6	677	2 D87018	probable asparagin
14	29	80.6	1184	1 A34795	kinase-related pr
15	28	77.8	163	2 B49233	S fimbrial adhesin
16	28	77.8	163	2 S15926	sfas protein precu
17	28	77.8	405	2 H71692	hypothetical prote
18	28	77.8	411	2 F97760	hypothetical prote
19	28	77.8	443	2 G64234	hypothetical prote
20	28	77.8	554	2 S67284	hypothetical prote
21	28	77.8	559	2 E89134	protein F25G6.4 [i
22	28	77.8	586	2 T24835	hypothetical prote
23	27	75.0	131	2 I40656	hypothetical prote
24	27	75.0	134	2 F85362	hypothetical prote
25	27	75.0	186	2 T50403	probable succinate
26	27	75.0	208	2 C71097	hypothetical prote
27	27	75.0	208	2 G97767	hypothetical prote
28	27	75.0	254	2 F97263	lactose phosphotra
29	27	75.0	288	2 A90135	SAM-dependent meth

30 27 75.0 311 2 A57223 pheromone receptor
31 27 75.0 328 2 B65086 hydrogenase (EC 1.
32 27 75.0 328 2 A85959 hydrogenase-2 smal
33 27 75.0 328 2 A91114 hydrogenase-2 smal
34 27 75.0 328 2 AE0885 hydrogenase-2 smal
35 27 75.0 337 1 WMBE82 ribonucleoside-dip
36 27 75.0 337 1 WMBE32 ribonucleoside-dip
37 27 75.0 340 1 WMBE57 ribonucleoside-dip
38 27 75.0 351 2 AH1702 ribonucleoside-dip
39 27 75.0 370 2 S95339 heat shock transcr
40 27 75.0 380 2 H90261 hypothetical prote
41 27 75.0 521 2 E64181 hypothetical prote
42 27 75.0 592 2 F81417 probable cytochrom
43 27 75.0 647 2 T28214 MCP-domain signal
44 27 75.0 669 2 E84852 probable nucleosid
45 27 75.0 697 2 T39512 hypothetical prote

ALIGNMENTS

RESULT 1

A71809

hypothetical protein jhp1413 - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 28-Jul-2000

C:Accession: A71809

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71809; MUID:99120557; PMID:9923682

A:Accession: A71809

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-115 <ARN>

A:Cross-references: GB:AE001564; GB:AE001439; NID:g4156032; PIDN:AAD06992.1; PID:g415

A:Experimental source: strain J99

C:Genetics:

A:Gene: jhp1413

C:Superfamily: Helicobacter pylori hypothetical protein HP1524

Query Match 83.3%; Score 30; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
DB 93 TNNVLQ 98

RESULT 2

D64710

hypothetical protein HP1524 - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 28-Jul-2000

C:Accession: D64710

R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKee

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: D64710

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-115 <TOM>

A:Cross-references: GB:AE000650; GB:AE000511; NID:g2314700; PIDN:AAD08566.1; PID:g231

C:Superfamily: Helicobacter pylori hypothetical protein HP1524

Query Match 83.3%; Score 30; DB 2; Length 115;

Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| | | | |
Db 93 TNNVLQ 98

RESULT 3

hypothetical protein Y18D10A.11 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T26523

R:Harris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z20226

A:Accession: T26523

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-248 <WIL>

A:Cross-references: EMBL:AL034393; PIDN:CAA22314.1; CESP:Y18D10A.11

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.11

A:Introns: 55/2; 101/2; 142/2; 204/3

Query Match 83.3%; Score 30; DB 2; Length 248;

Best Local Similarity 85.7%; Pred. No. 22;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8
| | | | |
Db 150 NNVLOXT 156

RESULT 4

sporulation protein SpoIID [imported] - Clostridium acetobutylicum

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001

C:Accession: A97252

R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: A97252

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-345 <KOR>

A:Cross-references: GB:AE001437; PIDN:AAK80804.1; PID:gl5025907; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC2861

C:Superfamily: stage II sporulation protein D

Query Match 83.3%; Score 30; DB 2; Length 345;

Best Local Similarity 85.7%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8
| | | | |
Db 271 NNVLOXT 277

RESULT 5

F81059

YhbXYjhw/Yajp/Yjdb family protein NMB1638 [imported] - Neisseria meningitidis (strain M

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001

C:Accession: F81059; E81816

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: F81059

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <DET>

A:Cross-references: GB:AE002514; GB:AE002098; NID:g7226886; PIDN:AAF41987.1; PID:g722

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: E81816

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <PAR>

A:Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85114.1; PID:g738

A:Experimental source: serogroup A, strain 22491

C:Genetics:

A:Gene: NMB1638; NMA1892

C:Superfamily: Escherichia coli yidB protein

Query Match 83.3%; Score 30; DB 2; Length 544;

Best Local Similarity 85.7%; Pred. No. 54;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8
| | | | |
Db 105 NNVLOXT 111

RESULT 6

T41941

glycoprotein B - human herpesvirus 7 (strain JI)

C:Species: human herpesvirus 7

A:Variety: strain JI

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T41941

R:Nicholas, J.

submitted to the EMBL Data Library, December 1995

A:Description: Determination and analysis of the complete nucleotide sequence of huma

A:Reference number: Z22022

A:Accession: T41941

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-822 <NIC>

A:Cross-references: EMBL:U43400; PIDN:AAC54701.1

A:Experimental source: strain JI

C:Genetics:

A:Note: U39

C:Superfamily: herpesvirus glycoprotein B

Query Match 83.3%; Score 30; DB 2; Length 822;

Best Local Similarity 75.0%; Pred. No. 87;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
| | | | |
Db 726 TNNVLQXT 733

RESULT 7

G84652

probable receptor-like protein kinase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84652

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y

M.; Koo, M.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: G84652

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-960 <STO>

A:Cross-references: GB:AE002093; NID:g3643604; PIDN:AAC42251.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g25790

A:Map position: 2

Query Match 83.3%; Score 30; DB 2; Length 960;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6
|||||

Db 932 TNNVLQ 937

RESULT 8

G87273

oxidoreductase iron/ascorbate family CC0200 [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: G87273

R: Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.F.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: G87273

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-311 <STO>

A:Cross-references: GB:AE005673; NID:gl3421323; PIDN:AAK22187.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0200

Query Match 80.6%; Score 29; DB 2; Length 311;

Best Local Similarity 75.0%; Pred. No. 49;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
|||||

Db 238 TNNVLQST 245

RESULT 9

A46352

ORF1 protein - *Chlorella virus PBCV-1*

C:Species: *Chlorella virus PBCV-1*

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Nov-2000

C:Accession: A46352; T18035

R: Schuster, A.M.; Graves, M.; Korth, K.; Ziegelbein, M.; Brumbaugh, J.; Grone, D.; Meint Virology 176, 515-523, 1990

A:Title: Transcription and sequence studies of a 4.3-kbp fragment from a ds-DNA eukaryotic virus.

A:Reference number: A46352; MUID:90266467; PMID:2345963

A:Accession: A46352

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-374 <SCH>

A:Cross-references: GB:M33758; NID:g323370; PIDN:AAA66400.1; PID:g807594

R: Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z18806

A:Accession: T18035

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-82, 'T', 84-374 <GRA>

A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96900.1

A:Experimental source: specific host *Chlorella* strain NC64A

C:Genetics:

A:Note: A533R

Query Match 80.6%; Score 29; DB 2; Length 374;

Best Local Similarity 85.7%; Pred. No. 61;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLQXT 8
|||||

Db 112 NNVLQXT 118

RESULT 10

T04199

hypothetical protein T4F9.100 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999

C:Accession: T04199

R: Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.

submitted to the Protein Sequence Database, March 1999

A:Reference number: Z15260

A:Accession: T04199

A:Molecule type: DNA

A:Residues: 1-407 <BEV>

A:Cross-references: EMBL:AL049523

A:Experimental source: cultivar Columbia; BAC clone T4F9

C:Genetics:

A:Map position: 4

A:Introns: 37/3; 114/3; 210/3; 251/3; 327/3

A:Note: T4F9.100

Query Match 80.6%; Score 29; DB 2; Length 407;

Best Local Similarity 75.0%; Pred. No. 67;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
|||||

Db 54 TNSVLQHT 61

RESULT 11

T43047

retrovirus-related env polyprotein homolog - Mediterranean fruit fly retrotransposon

C:Species: *Ceratitis capitata* (Mediterranean fruit fly)

C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-Jan-2000

C:Accession: T43047

R: Zhou, Q.; Haymer, D.S.

submitted to the EMBL Data Library, June 1996

A:Description: Gypsy-like retrotransposon in the Medfly.

A:Reference number: Z22299

A:Accession: T43047

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-501 <ZHO>

A:Cross-references: EMBL:U60529; NID:g1402846; PID:g1402849; PIDN:AAC28744.1

A:Experimental source: strain Med+; clone B

C:Genetics:

A:Mobile element: retrotransposon yoyo

Query Match 80.6%; Score 29; DB 2; Length 501;

Best Local Similarity 83.3%; Pred. No. 84;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6
|||||

Db 147 TNNILQ 152

RESULT 12

H84510

probable receptor-like protein kinase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84510
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84510
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-524 <STO>
A:Cross-references: GB:AE002093; NID:q4726119; PIDN:AAD28319.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g13800
A:Map position: 2

Query Match 80.6%; Score 29; DB 2; Length 524;
Best Local Similarity 83.3%; Pred. No. 89;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| | | | |
DB 44 TNNILQ 49

RESULT 13
D87018
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: D87018
R:Cole, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: D87018
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-677 <STO>
A:Cross-references: GB:AL450380; NID:g13092947; PIDN:CAC31255.1; GSPDB:GN00147
C:Genetics:
A:Gene: asnB

Query Match 80.6%; Score 29; DB 2; Length 677;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| | | | |
DB 39 TNNILQ 44

RESULT 14
A34795
Kinesin-related protein bimC - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 19-Jan-2001
C:Accession: A34795; B34795
R:Enos, A.P.; Morris, N.R.
Cell 60, 1019-1027, 1990
A:Title: Mutation of a gene that encodes a kinesin-like protein blocks nuclear division
A:Reference number: A34795; MUID:90199865; PMID:2138511
A:Accession: A34795
A:Molecule type: DNA
A:Residues: 1-1184 <ENO>
A:Cross-references: GB:M32075; NID:g168022; PIDN:AAA33298.1; PID:g168023
C:Superfamily: Kinesin-related protein Eg5; kinesin motor domain homology

C:Keywords: ATP; microtubule binding; mitosis; nucleotide binding; P-loop
F:82-422/Domain: kinesin motor domain homology <KNOT>
F:167-174/Region: nucleotide-binding motif A (P-loop)
F:173/Binding site: ATP (Lys) #status predicted
Query Match 80.6%; Score 29; DB 1; Length 1184;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 TNNVLQXT 8
| | | | |
DB 521 TNDVLOOT 528
RESULT 15
B49233
S fimbrial adhesin minor subunit sfas - Escherichia coli
C:Species: Escherichia coli
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1999
C:Accession: B49233
R:Hacker, J.; Kestler, H.; Hoshutzky, H.; Jann, K.; Lottspeich, F.; Korhonen, T.K.
Infect. Immun. 61, 544-550, 1993
A:Title: Cloning and characterization of the S fimbrial adhesin II complex of an Esch
A:Reference number: A49233; MUID:93138776; PMID:8093693
A:Contents: O18:K1
A:Accession: B49233
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-163 <HAC>
A:Cross-references: GB:S53210; NID:g264033; PIDN:AAB25046.1; PID:g264035
A:Note: sequence extracted from NCBI backbone (NCBIN:123181, NCBIP:123183)
C:Superfamily: conserved hypothetical protein bl503

Query Match 77.8%; Score 28; DB 2; Length 163;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
| | | | |
DB 30 TGNVLQRT 37

Search completed: February 26, 2003, 15:38:35
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:33:05 ; Search time 29 Seconds
(without alignments)
11.442 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	30	83.3	822	1 VGLB_HSV7J	P52352 human herpe
2	30	83.3	1762	1 DPOQ_HUMAN	O75417 homo sapien
3	29	80.6	1184	1 BINC_EMENI	P17120 emerichella
4	28	77.8	163	1 SPAS_ECOLI	P13430 escherichia
5	28	77.8	443	1 Y314_MFCGE	Q49415 mycoplasma
6	27	75.0	208	1 Y396_RICPR	Q92dd5 rickettsia
7	27	75.0	328	1 HYBA_ECOLI	P37179 escherichia
8	27	75.0	337	1 RIR2_HSV23	O92dd5 rickettsia
9	27	75.0	340	1 RIR2_HSV11	P03174 herpes simp
10	27	75.0	340	1 RIR2_HSV1K	P10224 herpes simp
11	27	75.0	365	1 SOXB_RHOSO	P06474 herpes simp
12	27	75.0	521	1 CYOA_HAEIN	P54997 rhodococcus
13	27	75.0	638	1 GOIM_COXBU	P45021 haemophilus
14	27	75.0	647	1 COAT_ADVG	P45650 coxiella bu
15	27	75.0	647	1 NTP1_MSEPV	P24029 aleutian mi
16	27	75.0	718	1 GUA_A_DICDI	O9Yw39 melanoplus
17	27	75.0	955	1 SVV_BUCAI	P32073 dictyosteli
18	27	75.0	993	1 DPOL_BPAPS	P57447 buchiera ap
19	27	75.0	2569	1 LMA3_MOUSE	Q9Tlq3 bacterioph
20	26	72.2	170	1 Y051_UREPA	O61789 mus musculu
21	26	72.2	274	1 TRV5_ANOGA	O9pr94 ureaplasma
22	26	72.2	283	1 YGZ0_YEAST	P35039 anopheles g
23	26	72.2	289	1 MWL1_YEAST	P53068 saccharomyc
24	26	72.2	324	1 HCRB_THAAR	P39731 saccharomyc
25	26	72.2	357	1 O2B2_HUMAN	O33820 thauera aro
26	26	72.2	423	1 YE18_SCHPO	O9g2k3 homo sapien
27	26	72.2	500	1 LCVB_TOBAC	O13873 schizosacch
28	26	72.2	522	1 MAL2_DROME	Q43578 nicotiana t
29	26	72.2	680	1 OPDA_ECOLI	P07190 drosophila
30	26	72.2	680	1 OPDA_SALTY	P27298 escherichia
31	26	72.2	681	1 OPDA_HAEIN	P27237 salmonella
32	26	72.2	734	1 UN36_CAEEL	P44573 haemophilus
33	26	72.2	750	1 YD33_MYCPN	P34374 caenorhabdi
					P75445 mycoplasma

34 26 72.2 804 1 VP5_WTV P12366 wound tumor
35 26 72.2 928 1 PM10_CHLPN Q9rb65 chlamydia p
36 26 72.2 1064 1 YY08_METJA Q60307 methanococ
37 26 72.2 1178 1 PHVB_SORBI P93527 sorghum bic
38 26 72.2 1205 1 YLJ9_CAEEL P34372 caenorhabdi
39 26 72.2 1233 1 YF16_YEAST P43597 caenorhabdi
40 26 72.2 1398 1 DNA2_SCHPO Q9ur02 schizosacch
41 25 69.4 68 1 Y244_RICPR O92dt1 rickettsia
42 25 69.4 69 1 GBG7_RAT P43425 rattus norv
43 25 69.4 109 1 RR10_CVACA O9tlv9 cyanidium c
44 25 69.4 113 1 DYLX_MOUSE P51807 mus musculu
45 25 69.4 148 1 Y222_TREPA O83251 treponema p

ALIGNMENTS

RESULT 1
VGLB_HSV7J
ID VGLB_HSV7J STANDARD; PRT; 822 AA.
AC P52352;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glycoprotein B precursor.
CN GB OR U39.
OS Human herpesvirus (type 7 / strain J1) (HHV7).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=57278;
RN [1]
RP SEQUENCE FROM N.A.
RA Nicholas J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN B FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; U43400; AAC54701.1; -;
InterPro: IPR000234; Glycoprot_B.
Pfam: PF00606; Glycoprotein_B; 1.
ProDom: PD000693; Glycoprot_B; 1.
Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 822 GLYCOPROTEIN B.
FT DOMAIN 23 649 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 650 670 POTENTIAL.
FT DOMAIN 671 683 CYTOPLASMIC (POTENTIAL).
FT TRANSEM 684 704 POTENTIAL.
FT DOMAIN 705 822 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 324 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 517 517 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 822 AA; 93147 MW; F2B410BD359C6498 CRC64;

Query Match 83.3%; Score 30; DB 1; Length 822;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TNNVLOXT 8
   ||:|||||
Db 726 TNSVLQAT 733

RESULT 2
DPOQ_HUMAN STANDARD; PRT; 1762 AA.
AC 075417; 095160;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase theta (BC 2.7.7.7) (DNA polymerase eta).
GN POLO OR POLH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA MEDLINE=89326527; PubMed=10395804;
RX Sharief F.S., Vojta P.J., Ropp P.A., Copeland W.C.;
RT "Cloning and chromosomal mapping of the human DNA polymerase theta
(POLO), the eighth human DNA polymerase.";
RN [2]
RL Genomics 59:90-96(1999).
CC
CC SEQUENCE OF 607-1762 FROM N.A.
CC Harris P.V., Kaelin C.B., Burtis K.C.;
CC "Catalytic activity of Pol eta, a new human DNA polymerase related to
CC the bacterial DNA polymerase I family and Drosophila Mus308.";
CC Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC
CC -!- FUNCTION: COULD BE INVOLVED IN THE REPAIR OF INTERSTRAND
CC CROSSLINKS.
CC
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-A FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF052573; AAC33565.1; -.
CC DR EMBL; AF043628; AAD05272.1; -.
CC DR HSSP; P00582; 1KFS.
CC DR Genew; HGNC:9186; POLO.
CC DR MIM; 604419;
CC DR InterPro; IPR001098; DNA_pol.
CC DR InterPro; IPR002298; DNA_polI.
CC DR Pfam; PF00476; DNA_pol_A; 1.
CC DR PRINTS; PR00868; DNAPOLI.
CC DR SMART; SM00482; POLAC; 1.
CC DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Transferase; DNA-directed DNA polymerase; ATP-binding; DNA repair.
FT NP_BIND 175 182 ATP (POTENTIAL).
FT CONFLICT 1185 1185 L -> F (IN REF. 2).
FT CONFLICT 1685 1685 R -> Q (IN REF. 2).
FT CONFLICT 1719 1719 V -> A (IN REF. 2).
FT CONFLICT 1756 1762 ELKDFDV -> RAKGL (IN REF. 2).
FT CONFLICT 1762 1762 ELKDFDV -> RAKGL (IN REF. 2).
SQ SEQUENCE 1762 AA; 197596 MW; AEC17A2103F6BDA CRC64;

Query Match 83.3%; Score 30; DB 1; Length 1762;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
   ||||||
Db 934 TNNVLQ 939

RESULT 3
BIMC_EMENI STANDARD; PRT; 1184 AA.
AC P17120;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Kinesin-like protein BIMC.
GN BIMC.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90199865; PubMed=2138511;
RA Enos A.P., Morris N.R.;
RT "Mutation of a gene that encodes a kinesin-like protein blocks
RT nuclear division in A. nidulans.";
RL Cell 60:1019-1027(1990).
CC -!- FUNCTION: IMPORTANT ROLE IN MITOTIC DIVIDING CELLS. MICROTUBULE
CC MOTOR REQUIRED FOR SPINDLE BODY SEPARATION.
CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. BIMC
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M32075; AAA33298.1; -.
CC DR PIR; A34795; A34795.
CC DR HSSP; P17119; 3KAR.
CC DR InterPro; IPR001752; kinesin_motor.
CC DR Pfam; PF00225; kinesin; 1.
CC DR PRINTS; PR00380; KINESINHEAVY.
CC DR SMART; SM00129; Kisc; 1.
CC DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
CC DR PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
KW Motor protein; Cell division; Microtubules; ATP-binding; Coiled coil;
KW Mitosis; Cell cycle; Phosphorylation.
FT DOMAIN 79 485 KINESIN-MOTOR (BY SIMILARITY).
FT DOMAIN 489 900 COILED COIL (POTENTIAL).
FT DOMAIN 901 1184 GLOBULAR (POTENTIAL).
FT NP_BIND 167 174 ATP (BY SIMILARITY).
FT MOD_RES 1006 1006 PHOSPHORYLATION (BY CDC2) (BY
FT SIMILARITY).
SQ SEQUENCE 1184 AA; 131630 MW; CEA015EC8F980E4F CRC64;

Query Match 80.6%; Score 29; DB 1; Length 1184;
Best Local Similarity 75.0%; Pred. No. 63;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
   ||:|||||
Db 521 TNDVLQQT 528

RESULT 4
SFAS_ECOLI STANDARD; PRT; 163 AA.
AC P13430;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-fimbrial adhesin protein sfas precursor.
GN SFAS.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=06:K15;
 RA MEDLINE=90138121; PubMed=2576095;
 RA Jann K., Hacker J.;
 RA "Analysis of genes coding for the static acid-binding adhesin and two
 RT other minor fimbrial subunits of the S-fimbrial adhesin determinant
 RT of Escherichia coli";
 RL Mol. Microbiol. 3:1735-1744(1989).
 RN [2]
 RP MOTAGENESIS OF LYS-138; ARG-140 AND LYS-144.
 RX MEDLINE=90307213; PubMed=2194961;
 RA Morschhauser J., Hoschuetzky H., Jann K., Hacker J.;
 RA "Functional analysis of the sialic acid-binding adhesin Sfas of
 RT pathogenic Escherichia coli by site-specific mutagenesis.";
 RL Infect. Immun. 58:2133-2138(1990).
 CC -!- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
 CC FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
 CC MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
 CC COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
 CC -!- FUNCTION: THIS PROTEIN IS THE SIALIC ACID-BINDING SUBUNIT OF THE
 CC S-FIMBRIAL ADHESINS. IT ALSO PLAYS A ROLE IN DETERMINATION OF THE
 CC DEGREE OF FIMBRATION OF THE CELL.
 CC -!- DISEASE: S-FIMBRIAL ADHESINS ENABLE PATHOGENIC E. COLI CAUSING
 CC URINARY-TRACT INFECTIONS OR NEWBORN MENINGITIS TO ATTACH TO
 CC GLYCOPROTEINS TERMINATING WITH ALPHA-SIALIC ACID- (2-3)-BETA-GAL.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X16664; CAA34653.1; .
 DR PIR: S06194; S06194.
 DR PIR: S15926; S15926.
 DR InterPro: IPR000259; Fimbrial.
 DR Pfam: PF00419; Fimbrial; 1.
 KW Fimbria; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 163
 FT DISULFID 38 75
 FT DOMAIN 138 144
 FT MUTAGEN 138 138
 FT MUTAGEN 140 140
 FT MUTAGEN 144 144
 FT MUTAGEN 144 144
 SQ SEQUENCE 163 AA; 17183 MW; 0BF1333BF9B2DE4F CRC64;
 Query Match 77.8%; Score 28; DB 1; Length 163;
 Best Local Similarity 75.0%; Pred. No. 10;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TNNVLOXT 8
 Db 30 TGNVLOXT 37
 RESULT 5
 Y314_MYCGE
 ID Y314_MYCGE STANDARD; PRT; 443 AA.
 AC Q49415; Q49279;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein MG314.
 GN MG314.

OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2097;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
 RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Boff K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium";
 RL Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE OF 70-171 FROM N.A.
 RC STRAIN=ATCC 33530 / G-37;
 RX MEDLINE=94075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Boff K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39712; AAC71536.1; .
 DR EMBL: U2151; AAC12432.1; .
 DR TIGR: MG314; .
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 443 AA; 51132 MW; 3EFB03F9168015D7 CRC64;
 Query Match 77.8%; Score 28; DB 1; Length 443;
 Best Local Similarity 71.4%; Pred. No. 34;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 NNVLQXT 8
 Db 158 NNILQOT 164
 RESULT 6
 Y396_RICPR
 ID Y396_RICPR STANDARD; PRT; 208 AA.
 AC Q9ZDD5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein RP396.
 GN RP396.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsia;
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria";
 RL Nature 396:133-140(1998).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AJ235271; CAAL4853.1; -
 DR InterPro: IPR000627; Dioxigenase.
 DR Pfam: PF00775; Dioxigenase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 208 AA; 23985 MW; 57BB82FADA62F864 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 208;
 Best Local Similarity 62.5%; Pred. No. 24;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 |||:|:|
 Db 48 TNNLLRKT 55

RESULT 7
 HYBA_ECOLI STANDARD; PRT; 328 AA.
 AC P37179;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hydrogenase-2 operon protein hybA precursor.
 GN HYBA OR B2996 OR Z4350 OR ECS3881.
 OS Escherichia coli, and
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_taxID=562, 83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / TGI;
 RX MEDLINE=94292472; PubMed=8021226;
 RA Menon N.K., Chatelus C.Y., Dervartanian M., Wendt J.C.,
 RA Shanmugam K.T., Peck H.D. Jr., Przybyla A.E.;
 RT "Cloning, sequencing, and mutational analysis of the hyb operon
 RT encoding Escherichia coli hydrogenase 2";
 RL J. Bacteriol. 176:4416-4423(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Godden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12";
 RL Science 277:1453-1474(1997).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533(2001).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,

RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 CC -!- FUNCTION: PARTICIPATES IN THE PERIPLASMIC ELECTRON-TRANSFERRING
 CC ACTIVITY OF HYDROGENASE 2 DURING ITS CATALYTIC TURNOVER.
 CC -!- COFACTOR: BINDS 3 4FE-4S CLUSTERS AND A 3FE-4S CLUSTER.
 CC -!- SUBCELLULAR LOCATION: Periplasmic.
 CC -!- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
 CC -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE THE SMALL SUBUNIT
 CC OF HYDROGENASE 2.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U09177; AAA21589.1; -
 DR EMBL: U28377; AAA59163.1; -
 DR EMBL: AE000382; AAC76032.1; -
 DR EMBL: AE003529; AAG58133.1; -
 DR EMBL: AF002563; BAB37304.1; -
 DR EcoGene; EGI1799; hybA.
 DR InterPro: IPR001450; 4Fe4s_ferredoxin.
 DR Pfam: PF00037; fer4; 1.
 DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
 KW Oxidoreductase; Signal; Periplasmic; Iron-sulfur; 4Fe-4S; 3Fe-4S;
 KW Complete proteome.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 328 HYDROGENASE-2 OPERON PROTEIN HYBA.
 FT METAL 47 47 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 50 50 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 53 53 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 57 57 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
 FT METAL 112 112 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).
 FT METAL 115 115 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).
 FT METAL 120 120 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).
 FT METAL 124 124 IRON-SULFUR 2 (3FE-4S) (POTENTIAL).
 FT METAL 145 145 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).
 FT METAL 148 148 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).
 FT METAL 151 151 IRON-SULFUR 3 (4FE-4S) (POTENTIAL).
 FT METAL 155 155 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
 FT METAL 174 174 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
 FT METAL 177 177 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
 FT METAL 193 193 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
 FT METAL 197 197 IRON-SULFUR 4 (4FE-4S) (POTENTIAL).
 SQ SEQUENCE 328 AA; 36003 MW; 77203A0F50F61662 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 328;
 Best Local Similarity 66.7%; Pred. No. 42;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||:|
 Db 83 TNNIIQ 88

RESULT 8
 RIR2_HSV23 STANDARD; PRT; 337 AA.
 ID RIR2_HSV23
 AC P03174;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
 DE (Ribonucleotide reductase) (38 kDa subunit).
 OS Herpes simplex virus (type 2 / strain 333).
 OS Herpes simplex virus (type 2 / strain HG52)
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

```

OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10313, 10315;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=333;
RX MEDLINE=84138764; PubMed=6321759;
RA Galloway D.A., Swain M.A.;
RT "Organization of the left-hand end of the herpes simplex virus type 2
  BgIII N fragment.";
RL J. Virol. 49:724-730(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=333;
RX MEDLINE=84057718; PubMed=6315408;
RA McLauchlan J., Clements J.B.;
RT "DNA sequence homology between two co-linear loci on the HSV genome
  which have different transforming abilities.";
RL EMBO J. 2:1953-1961(1983).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=HG52;
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
  thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
  thioresoxin.
CC -!- COFACTOR: BINDS 2 IRON IONS.
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
  SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M12700; AAA5807.1; -.
DR EMBL; X00048; CAA24930.1; -.
DR EMBL; Z86099; CAB06726.1; -.
DR PIR; A00528; WMBE32.
DR PIR; A00529; WMBE32.
DR InterPro: IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 91 91 IRON 1 (BY SIMILARITY).
FT METAL 121 121 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 124 124 IRON 1 (BY SIMILARITY).
FT METAL 184 184 IRON 2 (BY SIMILARITY).
FT METAL 218 218 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT ACT_SITE 128 128 IRON 2 (BY SIMILARITY).
FT CONFLICT 173 173 I -> V (IN REF. 2).
FT CONFLICT 235 235 G -> D (IN REF. 2).
SQ SEQUENCE 337 AA; 37625 MW; 1E27E96599EE2120 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 337;
Best Local Similarity 62.5%; Pred. No. 43;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
   |||:|:|
Db 200 TNNLLRVT 207

RESULT 9
RIR2_HSV11
ID RIR2_HSV11 STANDARD; PRT; 340 AA.

```

```

AC P10224;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-AUG-1990 (Rel. 15, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase) (38 kDa subunit).
GN UL40.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,
  RA McNaab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
  herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
  thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
  thioresoxin.
CC -!- COFACTOR: BINDS 2 IRON IONS.
CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
  SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D10879; BAA01686.1; -.
DR EMBL; X14112; CAA32303.1; -.
DR PIR; D30088; WMBE57.
DR InterPro: IPR000358; Ribonuc_redctse.
DR Pfam; PF00268; ribonuc_red_sm; 1.
DR PROSITE; PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 187 187 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 340 AA; 38019 MW; 4B4ED994BF74FD3F CRC64;

Query Match 75.0%; Score 27; DB 1; Length 340;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
   |||:|:|
Db 203 TNNLLRVT 210

RESULT 10
RIR2_HSV1K
ID RIR2_HSV1K STANDARD; PRT; 340 AA.
AC P06474;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase) (38 kDa subunit).
GN UL40.
OS Herpes simplex virus (type 1 / strain KOS).

```

OC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=10306;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=6292456;
 RX Draper K.G., Frink R.J., Wagner E.K.;
 RA Detailed characterization of an apparently unspliced beta herpes
 RT simplex virus type 1 gene mapping in the interior of another.";
 RL J. Virol. 43:1123-1128(1982).
 RN [2]
 RN REVISIONS.
 RA Wagner E.K.;
 RP Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RL
 CC -!- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
 CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
 CC thiorodoxin + H(2)O = ribonucleoside diphosphate + reduced
 CC thiorodoxin.
 CC
 CC -!- COFACTOR: BINDS 2 IRON IONS.
 CC
 CC -!- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 CC
 CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL CHAIN.
 CC
 CC -!- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
 CC SMALL CHAIN FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: J02212; AAA66436.1; -;
 DR InterPro: IPR000358; RibonucL_reductse.
 DR Pfam: PF00268; ribonuc_red_sm; 1.
 DR PROSITE: PS00368; RIBRED_SMALL; 1.
 KW Oxidoreductase; DNA replication; Iron.
 FT METAL 94 94 IRON 1 (BY SIMILARITY).
 FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
 FT METAL 127 127 IRON 1 (BY SIMILARITY).
 FT ACT_SITE 131 131 BY SIMILARITY.
 FT ACT_SITE 131 131 BY SIMILARITY.
 SQ SEQUENCE 340 AA; 37966 MW; 921DC04B9D278DE5 CRC64;

 Query Match 75.0%; Score 27; DB 1; Length 340;
 Best Local Similarity 62.5%; Pred. No. 44;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 TNNVLQXT 8
 DB 203 TNNLLRVT 210

 RESULT 11
 SOXB_RHOSO
 ID SOXB_RHOSO STANDARD; PRT; 365 AA.
 AC P54997;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Dibenzoethiophene desulfurization enzyme B (EC 3.1.2.24).
 GN SOXB OR DSZB.
 OS Rhodococcus sp. (strain IGTS8).
 OC Plasmid.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 OX NCBI_TaxID=1831;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=9505232; PubMed=7961424;
 RX Denome S.A., Oldfield C., Nash L.J., Young K.D.;
 RA "Characterization of the desulfurization genes from Rhodococcus sp.
 RT strain IGTS8.";
 RL J. Bacteriol. 176:6707-6716(1994).

RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96031556; PubMed=7574582;
 RX Piddington C.S., Kovacevich B.R., Rambosek J.;
 RA "Sequence and molecular characterization of a DNA region encoding the
 RT dibenzothiophene desulfurization operon of Rhodococcus sp. strain
 RT IGTS8.";
 RL Appl. Environ. Microbiol. 61:468-475(1995).
 CC -!- FUNCTION: PART OF A PATHWAY TO REMOVE COVALENTLY BOUND SULFUR FROM
 CC DIBENZOTHIOPHENE (DET) WITHOUT BREAKING CARBON-CARBON BONDS. THIS
 CC ENZYME METABOLIZES DET-SULFONE (DET-O2 OR DET 5,5-DIOXIDE) TO 2-
 CC HYDROXYBIPHENYL (2-HBP).
 CC -!- CATALYTIC ACTIVITY: 2-(2-hydroxyphenyl)benzenesulfinate + H(2)O =
 CC 2-hydroxybiphenyl + sulfite.
 CC -!- COFACTOR: FMN (POTENTIAL).
 CC -!- PATHWAY: SECOND STEP IN PATHWAY FROM DET TO 2-HBP.
 CC -!- SUBUNIT: HETERODIMER OF TWO SUBUNITS, SOXA AND SOXB.
 CC -!- SIMILARITY: BELONGS TO THE NTAA/SNAA/DSZA(SOXA) FAMILY OF
 CC MONOOXYGENASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U08850; AAA56672.1; -;
 DR EMBL: L37363; AAA99483.1; -;
 KW Hydrolase; Monooxygenase; Flavoprotein; FMN; Plasmid.
 SQ SEQUENCE 365 AA; 39044 MW; DA6A867756DA23D6 CRC64;

 Query Match 75.0%; Score 27; DB 1; Length 365;
 Best Local Similarity 83.3%; Pred. No. 47;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TNNVLQ 6
 DB 338 TNNLLQ 343

 RESULT 12
 CYOA_HAEIN
 ID CYOA_HAEIN STANDARD; PRT; 521 AA.
 AC P45021;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable cytochrome oxidase subunit I (EC 1.10.3.-).
 GN H11076.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: PROBABLE CYTOCHROME OXIDASE SUBUNIT.
 CC -!- COFACTOR: CONTAINS THE PROTOHEME IX CENTER B558 (BY SIMILARITY).

```

CC -!- SUBUNIT: HETERODIMER OF SUBUNITS I AND II (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (BY SIMILARITY).
CC -!- SIMILARITY: STRONG, TO BOTH E.COLI APPC AND CYDA. PROBABLE
CC ORTHOLOG OF THE ANCESTOR OF APPC/CYDA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32787; AAC22732.1; -
DR TIGR; H11076; -
DR InterPro; IPR002585; Bac_Ubq_Cox.
DR Pfam; PF01654; Bac_Ubq_Cox; 1.
DR Oxidoreductase; Electron transport; Transmembrane; Inner membrane;
KW Heme; Complete proteome.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 42 POTENTIAL.
FT DOMAIN 43 94 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 95 114 POTENTIAL.
FT DOMAIN 115 129 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 130 149 POTENTIAL.
FT DOMAIN 150 187 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 188 207 POTENTIAL.
FT DOMAIN 208 219 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 220 239 POTENTIAL.
FT DOMAIN 240 397 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 398 417 POTENTIAL.
FT DOMAIN 418 475 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 476 495 POTENTIAL.
FT DOMAIN 496 521 PERIPLASMIC (POTENTIAL).
FT METAL 186 186 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 521 AA; 58136 MW; FD44C20457D2265 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 521;
Best Local Similarity 62.5%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TNNVLQXT 8
Db 366 TNNVVDAT 373
||||:|

RESULT 13
60IM_COXBU STANDARD; PRT; 638 AA.
AC P45650;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 60 kDa inner-membrane protein homolog.
OS Coccidia burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coccidia group; Coccidia.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nine Mile phase I / Bratislava;
RX MEDLINE=94350801; PubMed=8071197;
RA Suhan M., Chen S.Y., Thompson H.A., Hoover T.A., Hill A.,
RA Williams J.C.;
RT "Cloning and characterization of an autonomous replication sequence
RT from Coccidia burnetii.";
RL J. Bacteriol. 176:5233-5243(1994).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Bairoch A.;
RL Unpublished observations (AUG-1995).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane

```

```

CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE OXAL / 60 KDA IMP FAMILY.
CC -!- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
CC INTRODUCED IN POSITION 482. THERE IS PROBABLY ANOTHER ERROR
CC IN THE C-TERMINAL PART AND THE REAL SEQUENCE COULD BE SHORTER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10529; AAA56919.1; ALT FRAME.
DR EMBL; U10529; AAA56921.1; ALT FRAME.
DR InterPro; IPR001708; 60kDa_innermemb.
DR Pfam; PF02096; 60KD_IMP; 1.
DR PRINTS; PR00701; 60KDINNERMP.
KW Transmembrane; Inner membrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 346 366 POTENTIAL.
FT TRANSMEM 369 389 POTENTIAL.
FT TRANSMEM 436 456 POTENTIAL.
SQ SEQUENCE 638 AA; 72903 MW; 6E670FC9EA2E80E2 CRC64;

Query Match 75.0%; Score 27; DB 1; Length 638;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TNNVLQ 6
Db 533 TNNVQ 538
||||:|

RESULT 14
COAT_ADVQ STANDARD; PRT; 647 AA.
AC P24029;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein VP1 [Contains: Coat protein VP2].
OS Aleutian mink disease parvovirus (strain G) (ADV).
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Parvovirus.
OX NCBI_TaxID=10783;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88275062; PubMed=2839709;
RA Bloom M.E., Alexandersen S., Perryman S., Lechner D.,
RA Wolfenbarger J.B.;
RT "Nucleotide sequence and genomic organization of Aleutian mink
RT disease parvovirus (ADV): sequence comparisons between a
RT nonpathogenic and a pathogenic strain of ADV.";
RL J. Virol. 62:2903-2915(1988).
CC -!- SIMILARITY: BELONGS TO THE PARVOVIRUSES COAT PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M20036; AAA66615.1; -
DR PIR; A35529; VCPVAP.
DR PIR; B36760; B36760.
DR HSSP; P30129; 4DPV.
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
KW Coat protein; Glycoprotein.
DOMAIN 22 39 POLY-GLY.

```

FT CARBOHYD 16 16 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 40 40 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 47 47 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 647 AA; 73517 MW; 22CE812094FFBFA CRC64;

Query Match 75.0%; Score 27; DB 1; Length 647;

Best Local Similarity 62.5%; Pred. No. 94;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8

Db 185 TNNILPYT 192

RESULT 15

NTPI_MSEPV STANDARD; PRT: 647 AA.
AC Q9YV39;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Nucleoside triphosphatase I (EC 3.6.1.15) (Nucleoside triphosphate
phosphohydrolase I) (NPH I).
GN MSV053.
OS Melanoplus sanguinipes entomopoxvirus (MsePV).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
OX NCBI_TaxID=83191;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Tucson;
RX MEDLINE=99102612; PubMed=9847359;
RA Afonso C.L., Tulman E.R., Lu Z., Oms E., Kutish G.F., Rock D.L.;
RT "The genome of Melanoplus sanguinipes entomopoxvirus";
RL J. Virol. 73:533-552(1999).
CC -!- FUNCTION: SERVES TWO ROLES IN TRANSCRIPTION; IT ACTS IN CONCERT
CC WITH VIRAL TERMINATION FACTOR/CAPPING ENZYME TO CATALYZE RELEASE
CC OF UUUUUU-CONTAINING NASCENT RNA FROM THE ELONGATION COMPLEX, AND
CC IT ACTS BY ITSELF AS A POLYMERASE ELONGATION FACTOR TO FACILITATE
CC READTHROUGH OF INTRINSIC PAUSE SITES (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.
CC -!- SIMILARITY: BELONGS TO THE NPH I SUBFAMILY OF HELICASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF063866; AAC37824.1; .
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR000330; SNF2_N.
DR Pfam; PF00176; SNF2_N; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00490; HELICC; 1.
KW Hydrolyase; ATP-binding; Transcription.
FT DOMAIN 40 322 SNF2_N.
FT DOMAIN 404 490 HELICASE_C.
FT NP_BIND 61 68 ATP (BY SIMILARITY).
FT SITE 150 153 DEXH BOX.
SQ SEQUENCE 647 AA; 75134 MW; E78C7F768235D51D CRC64;
Query Match 75.0%; Score 27; DB 1; Length 647;
Best Local Similarity 83.3%; Pred. No. 94;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
Db 197 TNNVLE 202

Search completed: February 26, 2003, 15:37:37
Job time : 30 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:33:36 : Search time 29 Seconds
(without alignments)
56.841 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriophage.*
- 17: sp_archaea.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	94.4	272	10 Q9ZTA9	Q9ZTA9 Dolichos la
2	34	94.4	279	10 Q9W7M4	Q9W7M4 phaseolus v
3	31	86.1	79	10 Q94AJ7	Q94AJ7 arabidopsis
4	30	83.3	115	16 Q26052	Q26052 helicobacte
5	30	83.3	115	16 Q9ZJA0	Q9ZJA0 helicobacte
6	30	83.3	248	5 Q9XW20	Q9XW20 caenorhabdi
7	30	83.3	345	16 Q97F80	Q97F80 clostridium
8	30	83.3	541	10 Q94LJ3	Q94LJ3 zea mays (m
9	30	83.3	544	16 Q34609	Q34609 neisseria m
10	30	83.3	822	12 Q39989	Q39989 human herpe
11	30	83.3	822	12 Q56280	Q56280 human herpe
12	30	83.3	822	12 Q96910	Q96910 human herpe
13	30	83.3	960	10 Q82318	Q82318 arabidopsis
14	30	83.3	1268	2 Q9RER7	Q9RER7 bacillus sp
15	30	83.3	2724	4 Q96SE4	Q96SE4 homo sapien
16	29	80.6	170	5 Q9NBF9	Q9NBF9 philodina r

17	29	80.6	284	10 Q9FYU9	Q9FYU9 sophora fla
18	29	80.6	311	16 Q9ABM7	Q9ABM7 caulobacter
19	29	80.6	374	12 Q66210	Q66210 paramecium
20	29	80.6	374	12 Q98583	Q98583 paramecium
21	29	80.6	393	10 Q9ZSC0	Q9ZSC0 arabidopsis
22	29	80.6	407	10 Q9T0C1	Q9T0C1 arabidopsis
23	29	80.6	501	5 Q17319	Q17319 ceratitidis c
24	29	80.6	524	10 Q9SKG4	Q9SKG4 arabidopsis
25	29	80.6	613	2 Q93LN4	Q93LN4 xanthomonas
26	29	80.6	649	12 Q91GN0	Q91GN0 epiphyas po
27	29	80.6	674	17 Q96XG2	Q96XG2 sulfolobus
28	29	80.6	677	16 Q9CCR2	Q9CCR2 mycobacteri
29	28	77.8	163	13 Q9IBS8	Q9IBS8 xenopus lae
30	28	77.8	163	2 Q53297	Q53297 escherichia
31	28	77.8	204	13 Q919D9	Q919D9 xenopus lae
32	28	77.8	212	15 Q9WGU7	Q9WGU7 human immun
33	28	77.8	351	5 Q966A1	Q966A1 caenorhabdi
34	28	77.8	363	16 Q8XRT3	Q8XRT3 raistonia s
35	28	77.8	390	16 Q8Y2H4	Q8Y2H4 raistonia s
36	28	77.8	405	16 Q9ZDH0	Q9ZDH0 rickettsia
37	28	77.8	411	16 Q92ID4	Q92ID4 rickettsia
38	28	77.8	466	5 Q22395	Q22395 caenorhabdi
39	28	77.8	554	3 Q08887	Q08887 saccharomyc
40	27	75.0	134	10 Q65553	Q65553 arabidopsis
41	27	75.0	186	3 Q9P7X0	Q9P7X0 schizosacch
42	27	75.0	200	2 Q8RTM2	Q8RTM2 helicobacte
43	27	75.0	208	16 Q92I77	Q92I77 rickettsia
44	27	75.0	229	12 Q9QCN4	Q9QCN4 aleutian mi
45	27	75.0	230	12 Q9QCN5	Q9QCN5 aleutian mi

ALIGNMENTS

RESULT 1

Q9ZTA9	ID	Q9ZTA9	PRELIMINARY;	PRT;	272 AA.
AC	Q9ZTA9;				
DT	01-MAY-1999 (TrEMBLrel. 10, Created)				
DT	01-MAY-1999 (TrEMBLrel. 10, Last sequence update)				
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)				
DE	Mannose lectin.				
GN	FRIL.				
OS	Dolichos lab lab (Field bean).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;				
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.				
OX	NCBI_TaxID=35936;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=COTYLEDON;				
RX	MEDLINE=99110944; PubMed=9892687;				
RA	Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;				
RT	"cDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves				
RT	hematopoietic progenitors in suspension culture."				
RL	Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).				
DR	EMBL; AF067417; AAD10734.1; -				
DR	HSSP; P02866; IONA.				
DR	InterPro; IPR000985; Lectin_lega.				
DR	InterPro; IPR001220; Lectin_legb.				
DR	Pfam; PF00138; lectin_lega; 1.				
DR	Pfam; PF00139; lectin_legb; 1.				
DR	ProDom; PD000671; Lectin_lega; 1.				
DR	ProDom; PD000711; Lectin_legb; 1.				
DR	PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.				
DR	PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.				
KW	Lectin.				
SQ	SEQUENCE 272 AA; 29900 MW; FA6C004307441495 CRC64;				

Query Match 94.4%; Score 34; DB 10; Length 272;
Best Local Similarity 87.5%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TNNVLOXT 8
DB 35 TNNVLOXT 42

RESULT 2
Q9M7M4 ID Q9M7M4 PRELIMINARY; PRT; 279 AA.
AC Q9M7M4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-WAR-2002 (TRENBLrel. 20, Last annotation update)
DE Mannose lectin FRIL (Fragment).
OS Phaseolus vulgaris (kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
RA Feldman M.;
RT "A new lectin in red kidney bean called PVFRL stimulates
RT proliferation of NIH3T3 cells expressing the Flt3 receptor.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121458; AAF28739.1;
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR ProDom; PD00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
FT NON_TER
FT SEQUENCE 279 AA; 31102 MW; F8919CF8B3E4652 CRC64;

Query Match 94.4%; Score 34; DB 10; Length 279;
Best Local Similarity 87.5%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 27 TNNVLOXT 34

RESULT 3
Q94AJ7 ID Q94AJ7 PRELIMINARY; PRT; 79 AA.
AC Q94AJ7;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Hypothetical 8.7 kDa protein.
GN AT4G33665 OR AT4G33666.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,
RA Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kaniya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesena E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene AT4G33665 (not previously annotated).";

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
RA Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
RA Southwick A., Davis R.W., Ecker J.R., Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY045998; AAK76672.1;
DR EMBL; AY079343; AAL85074.1;
KW Hypothetical protein.
SQ SEQUENCE 79 AA; 8660 MW; F8187D6D858360D0 CRC64;

Query Match 86.1%; Score 31; DB 10; Length 79;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 20 TNNVLOXT 27

RESULT 4
O26052 ID O26052 PRELIMINARY; PRT; 115 AA.
AC O26052;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-WAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein HPI524.
GN HPI524.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
DR EMBL; AE000650; AAD08566.1;
DR TIGR; HPI524;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 115 AA; 13287 MW; 1B56AA20E27EBE9D CRC64;

Query Match 83.3%; Score 30; DB 16; Length 115;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLO 6
DB 93 TNNVLO 98

RESULT 5
Q9ZJAO ID Q9ZJAO PRELIMINARY; PRT; 115 AA.
AC Q9ZJAO;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)

```

DE PUTATIVE.
 GN JHP1413.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Gull B.C., deJonghe B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori."
 RL Nature 397:176-180(1999).
 DR EMBL: AE001564; AAD06992.1; -.
 KW Complete proteome.
 SQ SEQUENCE 115 AA; 13299 MW; CD07B03053966DBF CRC64;

Query Match 83.3%; Score 30; DB 16; Length 115;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 Db 93 TNNVLQ 98
 |||||

RESULT 6
 Q9XW20
 ID Q9XW20 PRELIMINARY; PRT; 248 AA.
 AC Q9XW20;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Y18D10A.11 protein.
 GN Y18D10A.11.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Harris B.R.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=98519116;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RL Science 282:2012-2018(1998).
 DR EMBL: AL034393; CAA2314.1; -.
 SQ SEQUENCE 248 AA; 27963 MW; E7COA295197C6F6E CRC64;

Query Match 83.3%; Score 30; DB 5; Length 248;
 Best Local Similarity 85.7%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8
 Db 150 NNVLOAT 156
 |||||

RESULT 7
 Q97F80
 ID Q97F80 PRELIMINARY; PRT; 345 AA.
 AC Q97F80;
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)

DE Sporulation protein spoIID.
 GN CAC2861.
 OS Clostridium acetobutylicum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1488;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
 RA MEDLINE=21359325; PubMed=11466286;
 RA Noelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
 RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
 RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
 RA Bennett G.N., Koonin E.V., Smith D.R.;
 RT "Genome sequence and comparative analysis of the solvent-producing
 RT bacterium Clostridium acetobutylicum."
 RL J. Bacteriol. 183:4823-4838(2001).
 DR EMBL: AE007784; AAK80804.1; -.
 KW Complete proteome.
 SQ SEQUENCE 345 AA; 38653 MW; 2082986E5BDDFC3D CRC64;

Query Match 83.3%; Score 30; DB 16; Length 345;
 Best Local Similarity 85.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLOXT 8
 Db 271 NNVLOAT 277
 |||||

RESULT 8
 Q94IJ3
 ID Q94IJ3 PRELIMINARY; PRT; 541 AA.
 AC Q94IJ3;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Somatic embryogenesis receptor-like kinase 3 (Fragment).
 GN SERK3
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.
 OX NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. A188;
 RA Baudino S., Hansen S., Brettschneider R., Hecht V.F., Dresselhaus T.,
 RA Loerz H., Dumas C., Rogowsky P.M.;
 RT "Molecular characterization of two novel maize LRR receptor-like
 RT kinases, which belong to the SERK gene family."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ400870; CAC37642.1; -.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00560; LRR; 5.
 DR Pfam: PF00069; pkinase; 1.
 DR PRODOM: PD000001; Euk_pkinase; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 KW Kinase; Receptor.
 FT NON_TER 1
 FT NON_TER 541
 SQ SEQUENCE 541 AA; 59158 MW; FCC6BEA86BDF8587 CRC64;

Query Match 83.3%; Score 30; DB 10; Length 541;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||||


```

Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 726 TNSVLQAT 733
|||||

RESULT 12
Q96910
ID Q96910 PRELIMINARY; PRT; 822 AA.
AC Q96910;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Glycoprotein B.
GN U39.
OS Human herpesvirus 7.
OC Viruses; GSDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=10372;
RN [1]
RP SEQUENCE FROM N.A.
RA Hata A., Mukai T., Isegawa Y., Yamanishi K.;
RT "Identification and analyses of glycoprotein B of human herpesvirus
RT 7."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; D86577; BAAL3124.1; -.
DR InterPro; IPR000234; Glycoprot.B.
DR Pfam; PF00606; Glycoprotein_B; 1.
DR ProDom; PD000693; Glycoprot.B; 1.
SQ SEQUENCE 822 AA; 93178 MW; E3BFA858362A619 CRC64;

Query Match 83.3%; Score 30; DB 12; Length 822;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
DB 726 TNSVLQAT 733
|||||

RESULT 13
Q82318
ID Q82318 PRELIMINARY; PRT; 960 AA.
AC Q82318;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Putative receptor-like protein kinase.
GN AT2G25790.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;

```

```

Lin X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC005395; AAC42251.1; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003592; LRR_out.
DR Pfam; PF00560; LRR; 20.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00370; LRR; 16.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Kinase; Receptor; Transferase.
SQ SEQUENCE 960 AA; 106447 MW; 865D523C610DD838 CRC64;

Query Match 83.3%; Score 30; DB 10; Length 960;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
DB 932 TNNVLQ 937
|||||

RESULT 14
Q9RER7
ID Q9RER7 PRELIMINARY; PRT; 1268 AA.
AC Q9RER7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Surface layer protein precursor.
GN SBPA.
OS Bacillus sphaericus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1421;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 2177;
RA Iik N., Egelseer E.M., Jarosch M., Sleytr U.B., Sara M.;
RT "Nucleotide sequence of sbpA, the S-layer gene from Bacillus sphaericus
RT CCM 2177."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF211170; AAF22978.1; -.
DR HSP; P22629; 1SWL.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 3.
DR PROSITE; PS01072; SLH_DOMAIN; UNKNOWN_2.
KW Signal.
FT SIGNAL 1 30 POTENTIAL.
SQ SEQUENCE 1268 AA; 132046 MW; 2C51D40FADF0886 CRC64;

Query Match 83.3%; Score 30; DB 2; Length 1268;
Best Local Similarity 75.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
DB 513 TNNVLQAT 520
|||||

RESULT 15
Q96SE4
ID Q96SE4 PRELIMINARY; PRT; 2724 AA.
AC Q96SE4;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE DNA polymerase theta.
GN POLQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Abbas A.R., Linn S.M.;
 RT "Homo sapiens polymerase (DNA-directed), theta (POLQ), mRNA."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032677; AAK39635.1; -
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR001098; DNA_pol.
 DR InterPro; IPR001650; Helicase_C.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00476; DNA_pol_A; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR PROSITE; PS00447; DNA_POLYMERASE_A; UNKNOWN_1.
 KW ATP-binding; Helicase.
 SQ SEQUENCE 2724 AA; 305148 MW; 30B88663614E65DC CRC64;

Query Match 83.38; Score 30; DB 4; Length 2724;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||||
 Db 1896 TNNVLQ 1901

Search completed: February 26, 2003, 15:38:13
 Job time : 31 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:32:45 ; Search time 35 Seconds
(without alignments)
30.457 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	34	94.4	8	20 AAW87974
2	34	94.4	8	22 AAG62889
3	34	94.4	264	20 AAW87973
4	34	94.4	264	22 AAG62890
5	34	94.4	286	22 AAG62894
6	34	94.4	303	22 AAG62898
7	31	86.1	73	21 AAG59708
8	31	86.1	79	21 AAG59707
9	31	86.1	90	21 AAG59706
10	30	83.3	115	19 AAW98226

11	30	83.3	116	19	AAV11062	H. pylori ORF 06cp
12	30	83.3	544	23	AAU72915	Neisseria meningit
13	30	83.3	960	23	ABB91881	Herbicidally activ
14	29	80.6	524	21	AAAG32057	Arabidopsis thalia
15	29	80.6	524	23	AB91744	Herbicidally activ
16	29	80.6	579	23	AAO17257	A thaliana RKSIII
17	29	80.6	580	23	AAO17311	A thaliana recepto
18	29	80.6	612	21	AAAG51362	Arabidopsis thalia
19	29	80.6	1184	22	AAAG67415	Amino acid sequenc
20	27	75.0	25	23	ABG62350	Eubacterial DNA po
21	27	75.0	70	22	AAO09485	Human polypeptide
22	27	75.0	169	23	AAU47389	Rat pheromone rece
23	27	75.0	263	22	AAU15069	Protein encoded by
24	27	75.0	311	18	AAW19104	Rat pheromone rece
25	27	75.0	328	22	AAAG9806	E. coli growth and
26	27	75.0	337	19	AAW72207	HSV-2 strain SB5 C
27	27	75.0	337	19	AAW72099	HSV-2 strain SB5 C
28	27	75.0	337	19	AAW72019	HSV-2 strain SB5 C
29	27	75.0	365	15	AAW47874	Enzyme/biocatalyst
30	27	75.0	365	18	AAW09395	Rhodococcus rhodoc
31	27	75.0	365	20	AAW97239	An enzyme capable
32	27	75.0	365	20	AAW89242	Rhodococcus sp. st
33	27	75.0	365	22	AAU02430	Rhodococcus desulp
34	27	75.0	473	19	AAW61910	D. melanogaster gl
35	27	75.0	479	22	ABW60136	Drosophila melanog
36	27	75.0	632	23	ABP28314	Streptococcus poly
37	27	75.0	669	21	AAAG2215	Novel human diagno
38	27	75.0	705	22	ABG24213	C glutamicum prote
39	27	75.0	731	22	AAAG91100	Corynebacterium gl
40	27	75.0	731	22	AAW79423	Eucalyptus grandis
41	27	75.0	942	21	AAW25536	Shrimp white spot
42	27	75.0	1141	22	AAW85008	Human ORF2943 prot
43	26	72.2	63	23	ABP33970	Human foetal prote
44	26	72.2	68	22	AAW06379	Human ORFX protein
45	26	72.2	78	23	ABP00883	

ALIGNMENTS

RESULT 1
AAW87974

ID AAW87974 standard; Peptide; 8 AA.

AC AAW87974;

DT 13-APR-1999 (first entry)

DE Lectin derived progenitor cell preservation factor derived peptide.

KW Lectin derived progenitor cell preservation factor; progenitor cell;
KW haematopoietic cell; cultured cell preservation; anticancer therapy;
KW myeloblastic therapy; sickle-cell anaemia; ablative therapy protection;
KW FLK2/FLT3 receptor; ss.

OS Dolichos lab lab.

PN WO9859038-A1.

PD 30-DEC-1998.

PF 23-JUN-1998; 98WO-US13046.

PR 24-JUN-1997; 97US-0881189.

PA (IMCL-) IMCLONE SYSTEMS INC.

PA (REGC) UNIV CALIFORNIA.

PI Chrispeels MJ, Colucci MG, Moore JG;

DR WPI; 1999-081274/07.

PT New nucleic acid encoding plant lectin that preserves progenitor

PT cells - particularly haematopoietic progenitors, useful for bone
 PT marrow reconstitution after ablative therapy, and to increase DNA
 PT transfer in gene therapy

XX Claim 1; Page 46; 72pp; English.

XX The present sequence represents a peptide of lectin derived progenitor
 CC a cell preservation factor. The protein is used to preserve unipotent,
 CC pluripotent or totipotent progenitor cells, especially haematopoietic
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,
 CC kidney, liver, pancreas or thymus. Specific applications are
 CC preservation of cultured cells intended for administration after
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);
 CC treatment of tissues containing haematopoietic progenitors for subsequent
 CC transplant to improve haematopoietic competence; improving transfer of
 CC exogenous DNA to progenitor cells (in gene therapy of various
 CC haematological disorders, e.g. sickle-cell anaemia); and protection
 CC against ablative therapy (to eliminate proliferating cells specifically),
 CC followed by re-establishment of differentiation and proliferation of
 CC preserved progenitors. The protein, when linked to magnetic beads, may
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.

XX Sequence 8 AA;

Query Match 94.4%; Score 34; DB 20; Length 8;
 Best Local Similarity 87.5%; Pred. No. 7.8e+05;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 | | | | |
 Db 1 TNNVLQXT 8

RESULT 2

AAG62889
 ID AAG62889 standard; peptide; 8 AA.

XX AAG62889;

XX 17-SEP-2001 (first entry)

XX Peptide derived from a hyacinth bean FRIL polypeptide.

XX FRIL; FLK2/FLT3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Dolichos lab lab.

XX Key Location/Qualifiers

FT Misc-difference 7 /note= "not specified"

XX WO200149851-A1.

XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer

therapeutics -

XX Disclosure; Page 19; 173pp; English.

XX The present sequence is derived from a FRIL (FLK2/FLT3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of a
 CC therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 8 AA;

Query Match 94.4%; Score 34; DB 22; Length 8;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 | | | | |
 Db 1 TNNVLQXT 8

RESULT 3

AAW87973

ID AAW87973 standard; Protein; 264 AA.

XX AAW87973;

XX 13-APR-1999 (first entry)

XX A lectin derived progenitor cell preservation factor.

XX Lectin derived progenitor cell preservation factor; progenitor cell;
 KW haematopoietic cell; cultured cell preservation; anticancer therapy;
 KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;
 KW FLK2/FLT3 receptor.

XX Dolichos lab lab.

XX WO9859038-A1.

XX 30-DEC-1998.

XX 23-JUN-1998; 98WO-US13046.

XX 24-JUN-1997; 97US-0881189.

XX (IMCL-) IMCLONE SYSTEMS INC.
 XX (REGC) UNIV CALIFORNIA.

XX Chrispeels MJ, Colucci MG, Moore JG;

XX WPI; 1999-081274/07.

XX N-PSDB; AAX03593.

XX New nucleic acid encoding plant lectin that preserves progenitor
 PT cells - particularly haematopoietic progenitors, useful for bone
 PT marrow reconstitution after ablative therapy, and to increase DNA
 PT transfer in gene therapy

-Claim 1; Page 30-31; 72pp; English.

PS The present sequence represents a lectin derived progenitor cell
 CC preservation factor. The protein is used to preserve unipotent,
 CC pluripotent or totipotent progenitor cells, especially haematopoietic
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,
 CC kidney, liver, pancreas or thymus. Specific applications are
 CC preservation of cultured cells intended for administration after
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);
 CC treatment of tissues containing haematopoietic progenitors for subsequent
 CC transplant to improve haematopoietic competence; improving transfer of
 CC exogenous DNA to progenitor cells (in gene therapy of various
 CC haematological disorders, e.g. sickle-cell anaemia); and protection
 CC against ablative therapy (to eliminate proliferating cells specifically,
 CC followed by re-establishment of differentiation and proliferation of
 CC preserved progenitors. The protein, when linked to magnetic beads, may
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.

SQ Sequence 264 AA;

Query Match 94.4%; Score 34; DB 20; Length 264;
 Best Local Similarity 87.5%; Pred. No. 9.1;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TNNVLQXT 8
 |||||
 Db 27 TNNVLQVT 34

RESULT 4

AAG62890
 ID AAG62890 standard; Protein; 264 AA.

AC AAG62890;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a hyacinth bean FRIL polypeptide.

EX FRIL; FLK2/FLT3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Dolichos lab lab.

OS WO200149851-A1.

PN 12-JUL-2001.

PD 30-DEC-1999; 99WO-US31307.

PF 30-DEC-1999; 99WO-US31307.

PR (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42287.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics

XX Example 1; Page 54-55; 173pp; English.

XX The present sequence represents a FRIL (FLK2/FLT3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell

CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of
 CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 264 AA;

Query Match 94.4%; Score 34; DB 22; Length 264;
 Best Local Similarity 87.5%; Pred. No. 9.1;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TNNVLQXT 8

Db 27 TNNVLQVT 34

RESULT 5

AAG62894

XX AAG62894 standard; Protein; 286 AA.

AC AAG62894;

DT 17-SEP-2001 (first entry)

DE Alpha-amylase inhibitor signal peptide and FRIL fusion.

EX FRIL; FLK2/FLT3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair;
 KW alpha-amylase inhibitor gene.

OS Synthetic

OS Unidentified.

OS Dolichos lab lab.

XX WO200149851-A1.

PN 12-JUL-2001.

PD 30-DEC-1999; 99WO-US31307.

PF 30-DEC-1999; 99WO-US31307.

PR (PHYL-) PHYLOGIX LLC.

PI Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42295.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics

XX Example 1; Page 59; 173pp; English.

XX The present sequence represents fusion protein of alpha-amylase inhibitor
 CC signal peptide and FRIL (FLK2/FLT3 tyrosine kinase receptor-interacting
 CC lectin). The specification describes a composition of one or more members

CC of FRIL family of progenitor cell preservation factors. The composition
 CC is useful for alleviating or reducing the hematopoietic progenitor
 CC cell-depleting activity of a therapeutic treatment, including
 CC radiotherapeutic and/or chemotherapeutic treatment. Administration of
 CC FRIL compositions to a patient prior to treatment of the patient with
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting
 CC activity alleviates or reduces the hematopoietic progenitor
 CC cell-depleting activity of the therapeutic treatment in the patient.
 CC FRIL family members are useful for isolating population of progenitor
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is
 CC administered to reduce progenitor cell depleting effects of
 CC chemotherapeutic, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX
 XX
 SQ Sequence 286 AA;
 Query Match 94.4%; Score 34; DB 22; Length 286;
 Best Local Similarity 87.5%; Pred No. 10;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 DB 49 TNNVLQVT 56
 |||||

RESULT 6
 AAG62898
 ID AAG62898 standard; Protein: 303 AA.

XX
 AC AAG62898;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of a french bean FRIL polypeptide.

KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Phaseolus vulgaris.

XX WO200149851-A1.

PN 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

DR N-PSDB; AAH42306.

XX Legume progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -

XX Example 5; Page 81; 173pp; English.

XX The present sequence represents a FRIL (Flk2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of

CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of the chemotherapeutic, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX
 XX
 SQ Sequence 303 AA;

Query Match 94.4%; Score 34; DB 22; Length 303;
 Best Local Similarity 87.5%; Pred No. 11;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 DB 27 TNNVLQLT 34
 |||||

RESULT 7

AAG59708

ID AAG59708 standard; Protein: 73 AA.

XX
 AC AAG59708;

XX 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 77260.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

XX 05-MAR-1999; 99US-0123180.

XX 09-MAR-1999; 99US-0123548.

XX 23-MAR-1999; 99US-0125788.

XX 25-MAR-1999; 99US-0126264.

XX 29-MAR-1999; 99US-0126785.

XX 01-APR-1999; 99US-0127462.

XX 06-APR-1999; 99US-0128234.

XX 08-APR-1999; 99US-0128714.

XX 16-APR-1999; 99US-0129845.

XX 19-APR-1999; 99US-0130077.

XX 21-APR-1999; 99US-0130449.

XX 23-APR-1999; 99US-0130510.

XX 23-APR-1999; 99US-0130891.

XX 28-APR-1999; 99US-0131449.

XX 30-APR-1999; 99US-0132048.

XX 30-APR-1999; 99US-0132407.

XX 04-MAY-1999; 99US-0132484.

XX 05-MAY-1999; 99US-0132485.

XX 06-MAY-1999; 99US-0132486.

XX 06-MAY-1999; 99US-0132487.

XX 07-MAY-1999; 99US-0132863.

XX 11-MAY-1999; 99US-0134256.

XX 14-MAY-1999; 99US-0134218.

XX 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140595.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 16-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160758.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

```
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 73;
Best Local Similarity 75.0%; Pred. No. 9.6;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TNNVLOXT 8
Db 14 TNNVLOXT 21

RESULT 8
AAG59707
ID AAG59707 standard; Protein; 79 AA.
XX AC AAG59707;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 77259.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 06-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
```

PR	53-AUG-1999;	99US-0147030;
PR	04-AUG-1999;	99US-0147204;
PR	04-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147305;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148341;
PR	13-AUG-1999;	99US-0148565;
PR	13-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149175;
PR	18-AUG-1999;	99US-0149426;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	23-AUG-1999;	99US-0149929;
PR	23-AUG-1999;	99US-0149902;
PR	23-AUG-1999;	99US-0149930;
PR	25-AUG-1999;	99US-0150566;
PR	26-AUG-1999;	99US-0150884;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	30-AUG-1999;	99US-0151080;
PR	30-AUG-1999;	99US-0151303;
PR	31-SEP-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0152363;
PR	13-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0153758;
PR	16-SEP-1999;	99US-0154018;
PR	20-SEP-1999;	99US-0154039;
PR	22-SEP-1999;	99US-0154779;
PR	23-SEP-1999;	99US-0155139;
PR	24-SEP-1999;	99US-01555486;
PR	28-SEP-1999;	99US-01556559;
PR	29-SEP-1999;	99US-0156458;
PR	04-OCT-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	06-OCT-1999;	99US-0157753;
PR	07-OCT-1999;	99US-0157865;
PR	08-OCT-1999;	99US-0158029;
PR	12-OCT-1999;	99US-0158232;
PR	13-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	22-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	23-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161990;

PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity		
Matches		
6; Conservative		
0; Mismatches		
2; Indels		
Qy	1	TNNVLOXT 8
Db	20	TNNTLOTT 27
RESULT 9		
AAG59706		
ID	AAG59706 standard; Protein; 90 AA.	
AC	AAG59706;	
DT	18-OCT-2000 (first entry)	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 77258.	
XX	Protein identification; signal transduction pathway; metabolic	
KW	hybridisation assay; genetic mapping; gene expression control	
KW	termination sequence.	
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX	EPI033405-A2.	
PN	06-SEP-2000.	
PD	25-FEB-2000; 2000EP-0301439.	
XX	25-FEB-1999;	
XX	99US-0121825.	
PR	05-MAR-1999;	
PR	99US-0123180.	
PR	09-MAR-1999;	
PR	99US-0123548.	
PR	23-MAR-1999;	
PR	99US-0125788.	
PR	25-MAR-1999;	
PR	99US-0126264.	
PR	29-MAR-1999;	
PR	99US-0126785.	
PR	01-APR-1999;	
PR	99US-0127462.	
PR	06-APR-1999;	
PR	99US-0128234.	
PR	18-APR-1999;	
PR	99US-0128714.	
PR	06-APR-1999;	
PR	99US-0129845.	
PR	19-APR-1999;	
PR	99US-0130077.	
PR	21-APR-1999;	
PR	99US-0130449.	
PR	23-APR-1999;	
PR	99US-0130510.	
PR	23-APR-1999;	
PR	99US-0130891.	
PR	28-APR-1999;	
PR	99US-0131449.	
PR	30-APR-1999;	
PR	99US-0132048.	
PR	30-APR-1999;	
PR	99US-0132407.	
PR	04-MAY-1999;	
PR	99US-0132484.	
PR	05-MAY-1999;	
PR	99US-0132485.	
PR	06-MAY-1999;	
PR	99US-0132486.	
PR	07-MAY-1999;	
PR	99US-0132487.	
PR	11-MAY-1999;	
PR	99US-0132863.	
PR	14-MAY-1999;	
PR	99US-0134207.	
PR	14-MAY-1999;	
PR	99US-0134218.	
PR	14-MAY-1999;	
PR	99US-0134219.	
PR	14-MAY-1999;	
PR	99US-0134221.	
PR	14-MAY-1999;	
PR	99US-0134370.	
PR	18-MAY-1999;	
PR	99US-0134768.	
PR	19-MAY-1999;	
PR	99US-0134941.	
PR	20-MAY-1999;	
PR	99US-0135124.	
PR	21-MAY-1999;	
PR	99US-0135353.	
PR	24-MAY-1999;	
PR	99US-0135629.	
PR	25-MAY-1999;	
PR	99US-0136021.	
PR	27-MAY-1999;	
PR	99US-0136392.	
PR	28-MAY-1999;	
PR	99US-0136782.	
PR	01-JUN-1999;	
PR	99US-0137222.	
PR		

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140333.
PR 23-JUN-1999; 99US-0140334.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140901.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142134.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 86.1%; Score 31; DB 21; Length 90;
Best Local Similarity 75.0%; Pred. No. 12;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 DB 31 TNNFLQTT 38

RESULT 10
 AAW98226
 ID AAW98226 standard; Protein; 115 AA.
 XX AC
 XX AAW98226;
 XX 31-MAR-1999 (first entry)
 XX H. pylori GHPO 1099 protein.
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 KW peptic ulcer disease.
 XX Helicobacter pylori.
 XX OS
 XX PN WO9843478-A1.
 XX PD 08-OCT-1998.
 XX PF 01-APR-1998; 98WO-US06371.
 XX PR 29-JUL-1997; 97US-0902615.
 XX PR 01-APR-1997; 97US-0833457.
 XX PR 24-JUN-1997; 97US-0881227.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX
 PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
 XX WPI; 1998-542293/46.
 DR N-PSDB; AAX13945.
 XX
 PT New isolated Helicobacter polynucleotides - used to develop products
 PT for the diagnosis, prevention and treatment of Helicobacter
 PT infections and gastrointestinal diseases
 XX
 PS Claim 8; Page 159-160; 2054pp; English.
 CC This sequence represents a Helicobacter pylori GHPO protein of the
 CC invention. The polypeptides can be used for preventing or treating
 CC Helicobacter infections, and gastroduodenal diseases associated with
 CC these infections, including acute, chronic, and atrophic gastritis, and
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
 CC used for the production of antibodies. The products can also be used for
 CC detection and diagnosis.
 XX
 SQ Sequence 115 AA;
 Query Match 83.3%; Score 30; DB 19; Length 115;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TNNVLQ 6
 DB 93 TNNVLQ 98
 IIIII

RESULT 11
 AAY11062
 ID AAY11062 standard; Protein; 116 AA.
 XX AC
 XX AAY11062;
 XX 08-JUN-1999 (first entry)
 XX H. pylori ORF 06cp30603_10744075_c3_136 secreted protein.

KW Vaccine; probe; diagnostic; ORF; cell envelope protein;
 KW secreted protein; cytoplasmic protein; cellular protein.
 XX Helicobacter pylori.
 XX OS
 XX PN WO9824475-A1.
 XX PD 11-JUN-1998.
 XX PF 05-DEC-1997; 97WO-US22104.
 XX PR 14-JUL-1997; 97US-0891928.
 XX PR 05-DEC-1996; 96US-0759625.
 XX PR 25-MAR-1997; 97US-0823745.
 XX (ASTR) ASTRA AB.
 XX PI Alm RA, Castriotta LM, Doig PC, Kabok Z, Smith D;
 XX WPI; 1998-333051/29.
 DR N-PSDB; AAX30591.
 XX New isolated Helicobacter pylori nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of infection by
 PT H. pylori and other Helicobacter species
 XX
 PS Claims 37, 41; Page 224-225; 339pp; English.
 XX Recombinant or substantially pure preparations of H. pylori polypeptides
 CC are disclosed, together with the nucleic acids encoding them. In all,
 CC 97 ORFs are shown. The proteins are variously cell envelope proteins,
 CC cytoplasmic proteins, secreted proteins or other cellular proteins.
 CC Vaccines containing the nucleic acids or proteins are claimed, as are
 CC probes containing at least 8 nucleotides from the nucleic acid
 CC sequences. The vaccines are useful for treating or reducing the risk of
 CC H. pylori infections, and the probes can be used diagnostically for
 CC detecting the presence of Helicobacter in a sample. The products are
 CC also of use in screening for compounds having the ability to interfere
 CC with the H. pylori life cycle or to inhibit H. pylori infection.
 XX
 SQ Sequence 116 AA;
 Query Match 83.3%; Score 30; DB 19; Length 116;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TNNVLQ 6
 DB 94 TNNVLQ 99
 IIIII

RESULT 12
 AAU72915
 ID AAU72915 standard; Protein; 544 AA.
 XX AC
 XX AAU72915;
 XX 12-MAR-2002 (first entry)
 XX Neisseria meningitidis virulence protein #5.
 XX Meningitis; virulence; gene; antibacterial; vaccine; veterinary;
 KW infection; Gram-negative bacteria; antimicrobial.
 XX Neisseria meningitidis.
 XX OS
 XX PN WO200185772-A2.
 XX PD 15-NOV-2001.
 XX PF 08-MAY-2001; 2001WO-GB02003.
 XX PR 08-MAY-2000; 2000GB-0011108.

XX PA (MICR-) MICROSCIENCE LTD.
XX PI Tang C;
XX DR WPI; 2002-066593/09.
XX DR N-PSDB; AAS97200.
XX PT New peptide encoded by operon including virulence genes of *Neisseria*
XX PT meningitidis, useful as vaccine component for treating or preventing
XX PT meningitis and for identifying antimicrobial drug
XX PS Claim 4; Page 37-39; 423pp; English.
XX CC The invention relates to a peptide (I) encoded by an operon (II) of
XX CC *Neisseria meningitidis* including virulence genes, or a related molecule
XX CC having a 40% sequence similarity at the peptide or nucleotide level in a
XX CC Gram-negative bacterium, or its functional fragment, for therapeutic or
XX CC diagnostic use. (I) and (II) are useful in the manufacture of a
XX CC medicament for treating or preventing a condition (e.g., meningitis)
XX CC associated with infection by *Neisseria* or Gram-negative bacteria. The
XX CC product is useful for veterinary treatment and in a screening assay for
XX CC the identification of an antimicrobial drug. The vaccines have
XX CC prophylactic applications. AAU72911-AAU73014 represent *N. meningitidis*
XX CC virulence proteins of the invention.
XX SQ Sequence 544 AA;
Query Match 83.3%; Score 30; DB 23; Length 544;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 NNVLQXT 8
DB 105 NNVLQTT 111
RESULT 13
ABB91881
ID ABB91881 standard; Protein; 960 AA.
XX AC ABB91881;
XX DT 31-MAY-2002 (first entry)
XX DE Herbicidally active polypeptide SEQ ID NO 1092.
XX KW Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX PN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EF09892.
XX PR 28-AUG-2001; 2001WO-EF09892.
XX PA (FARB) BAYER AG.
XX PI Tietjen K, Weidler M;
XX DR WPI; 2002-269010/31.
XX PT Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms -
XX PS Claim 5; SEQ ID NO 1092; 261pp + Sequence Listing; English.
XX CC The invention relates to identifying target proteins

CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX Sequence 960 AA;
SQ Query Match 83.3%; Score 30; DB 23; Length 960;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNNVLQ 6
DB 932 TNNVLQ 937
RESULT 14
AAG32057
ID AAG32057 standard; Protein; 524 AA.
XX AC AAG32057;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38603.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127462.
XX PR 06-APR-1999; 99US-0128234.
XX PR 18-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 28-APR-1999; 99US-0130891.
XX PR 30-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 07-MAY-1999; 99US-0132863.
XX PR 11-MAY-1999; 99US-0134256.
XX PR 14-MAY-1999; 99US-0134216.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140695.
PR 29-JUN-1999; 99US-0140823.
PR 30-JUN-1999; 99US-0140991.
PR 01-JUL-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 80.6%; Score 29; DB 21; Length 524;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||:|
 Db 44 TNNILQ 49

RESULT 15

ABB91744
 ID ABB91744 standard; Protein; 524 AA.

XX ABB91744;

XX 31-MAY-2002 (first entry)

XX Herbicidally active polypeptide SEQ ID NO 955.

XX Herbicidal; plant; agriculture; herbicide.

XX Arabidopsis thaliana.

XX WO200210210-A2.

XX 07-FEB-2002.

XX 28-AUG-2001; 2001WO-EP09892.

XX 28-AUG-2001; 2001WO-EP09892.

XX (FARB) BAYER AG.

XX Tietjen K, Weidler M;

XX WPI; 2002-269010/31.

XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -

XX Claim 5; SEQ ID NO 955; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

XX Sequence 524 AA;

Query Match 80.6%; Score 29; DB 23; Length 524;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||:|
 Db 44 TNNILQ 49

Search completed: February 26, 2003, 15:37:01
 Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:38:16 ; Search time 32 Seconds
(without alignments)
9.430 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications-AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	83.3	115	10	US-09-881-752A-52
2	27	75.0	311	10	US-09-898-416-10
3	27	75.0	731	9	US-09-738-626-4854
4	26	72.2	234	9	US-09-764-868-1025
5	26	72.2	500	10	US-09-323-998D-59
6	26	72.2	555	9	US-09-764-868-878
7	25	69.4	181	10	US-09-864-761-34751
8	25	69.4	166	10	US-09-760-541-6
9	25	69.4	252	9	US-09-738-626-4420
10	25	69.4	262	9	US-10-260-877-2
11	25	69.4	262	10	US-09-815-242-10954
12	25	69.4	308	10	US-09-925-299-930
13	25	69.4	500	10	US-09-765-272-122
14	25	69.4	522	10	US-09-965-313-7
15	25	69.4	546	9	US-09-905-291A-250
16	25	69.4	546	9	US-09-902-853-250
17	25	69.4	546	9	US-09-907-824-250
18	25	69.4	546	9	US-09-907-841-250
19	25	69.4	546	9	US-09-904-011-250

20	25	69.4	546	9	US-10-174-590-412	Sequence 412, App
21	25	69.4	546	9	US-10-176-758-412	Sequence 412, App
22	25	69.4	546	9	US-10-175-737-412	Sequence 412, App
23	25	69.4	546	9	US-09-906-742-250	Sequence 250, App
24	25	69.4	546	9	US-10-173-706-412	Sequence 412, App
25	25	69.4	546	9	US-10-175-738-412	Sequence 412, App
26	25	69.4	546	9	US-10-175-752-412	Sequence 412, App
27	25	69.4	546	9	US-10-176-482-412	Sequence 412, App
28	25	69.4	546	9	US-10-176-757-412	Sequence 412, App
29	25	69.4	546	9	US-10-176-913-412	Sequence 412, App
30	25	69.4	546	9	US-10-180-552-412	Sequence 412, App
31	25	69.4	546	9	US-10-180-557-412	Sequence 412, App
32	25	69.4	546	9	US-09-906-838-250	Sequence 250, App
33	25	69.4	546	9	US-09-907-613-250	Sequence 250, App
34	25	69.4	546	9	US-09-907-942-250	Sequence 250, App
35	25	69.4	546	9	US-10-173-700-412	Sequence 412, App
36	25	69.4	546	9	US-10-174-572-412	Sequence 412, App
37	25	69.4	546	9	US-10-174-579-412	Sequence 412, App
38	25	69.4	546	9	US-10-174-582-412	Sequence 412, App
39	25	69.4	546	9	US-10-174-588-412	Sequence 412, App
40	25	69.4	546	9	US-10-175-739-412	Sequence 412, App
41	25	69.4	546	9	US-10-175-740-412	Sequence 412, App
42	25	69.4	546	9	US-10-175-743-412	Sequence 412, App
43	25	69.4	546	9	US-10-176-488-412	Sequence 412, App
44	25	69.4	546	9	US-10-176-492-412	Sequence 412, App
45	25	69.4	546	9	US-10-176-747-412	Sequence 412, App

ALIGNMENTS

RESULT 1
US-09-881-752A-52
; Sequence 52, Application US/09881752A
; Patent No. US20020115078A1

; GENERAL INFORMATION:

; APPLICANT: Kleanthous, Harold

; APPLICANT: Al-Garawi, Amal

; APPLICANT: Miller, Charles

; APPLICANT: Tomb, Jean-Francois

; APPLICANT: Oomen, Raymond P.

; TITLE OF INVENTION: Identification of Polynucleotides

; TITLE OF INVENTION: Encoding No. US20020115078A1 Helicobacter Polypeptides in t

; FILE REFERENCE: 06132/041002

; CURRENT APPLICATION NUMBER: US/09/881,752A

; CURRENT FILING DATE: 2001-06-15

; PRIOR APPLICATION NUMBER: US 08/833,457

; PRIOR FILING DATE: 1997-04-01

; NUMBER OF SEQ ID NOS: 370

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 52

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Helicobacter pylori

US-09-881-752A-52

Query Match 83.3%; Score 30; DB 10; Length 115;

Best Local Similarity 100.0%; Pred. No. 6.9;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6

Db 93 TNNVLQ 98

RESULT 2

US-09-898-416-10

; Sequence 10, Application US/09898416

; Patent No. US20020076759A1

; GENERAL INFORMATION:

; APPLICANT: Dulac, Catherine

; APPLICANT: Axel, Richard

```
; TITLE OF INVENTION: Cloning Of Vertebrate Pheromone Receptors And Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 0575/48557A/JPW/ADM
; CURRENT APPLICATION NUMBER: US/09/898,416
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 08/731,745
; PRIOR FILING DATE: 1996-10-18
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 311
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-898-416-10

Query Match          75.0%; Score 27; DB 10; Length 311;
Best Local Similarity 62.5%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 163 TNNFMQVT 170

RESULT 3
US-09-738-626-4854
; Sequence 4854, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4854
; LENGTH: 731
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4854

Query Match          75.0%; Score 27; DB 9; Length 731;
Best Local Similarity 62.5%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 639 TNNILAF 646

RESULT 4
US-09-764-868-1025
; Sequence 1025, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
```

```
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1025
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (5)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-1025

Query Match          72.2%; Score 26; DB 9; Length 234;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQ 6
Db 203 SNNVLQ 208

RESULT 5
US-09-323-998D-59
; Sequence 59, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-05019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Nicotiana sp.
US-09-323-998D-59

Query Match          72.2%; Score 26; DB 10; Length 500;
Best Local Similarity 71.4%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLQXT 8
Db 493 NNLLQDT 499

RESULT 6
US-09-764-868-878
; Sequence 878, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
```

; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 878
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-878

Query Match 72.2%; Score 26; DB 9; Length 555;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
Db 524 SNNVLQ 529

RESULT 7
US-09-864-761-34751
; Sequence 34751, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34751

LENGTH: 81
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000352.2
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.3
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.9
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.7
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 9.9
OTHER INFORMATION: SWISSPROT HIT: Q9Y6J0, EVALUE 2.00e-39
OTHER INFORMATION: EST_HUMAN HIT: A0120832.1, EVALUE 4.80e+00
US-09-864-761-34751

Query Match 59.4%; Score 25; DB 10; Length 81;
Best Local Similarity 66.7%; Pred. No. 59;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
Db 76 TNNLIQ 81

RESULT 8
US-09-760-541-6
; Sequence 6, Application US/09760541
; Patent No. US20010019713A1
; GENERAL INFORMATION:
; APPLICANT: Blazar, Beverly A.
; APPLICANT: Webb, Andrew C.
; TITLE OF INVENTION: NO. US20010019713A1 Interleukin Compositions and Methods
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/760,541
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,387
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: BLAZ-1.C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-760-541-6

```
Query Match          69.4%; Score 25; DB 10; Length 166;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 NNVLQXT 8
Db 55 NNLSST 61

RESULT 9
US-09-738-626-4420
; Sequence 4420, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4420
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4420

Query Match          69.4%; Score 25; DB 9; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 223 TNNVL 227

RESULT 10
US-10-260-877-2
; Sequence 2, Application US/10260877
; Publication No. US20030021813A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Chovan, Linda E.
; APPLICANT: Hessler, Paul E.
; APPLICANT: Reich, Karl A.
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF
; FILE REFERENCE: 6565.US.P1
; CURRENT APPLICATION NUMBER: US/10/260,877
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US/09/649,145
; PRIOR FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 262
; TYPE: PRT
```

```
; ORGANISM: H. influenzae
US-10-260-877-2

Query Match          69.4%; Score 25; DB 9; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 59 TNNVL 63

RESULT 11
US-09-815-242-10954
; Sequence 10954, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10954
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-09-815-242-10954

Query Match          69.4%; Score 25; DB 10; Length 262;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 59 TNNVL 63

RESULT 12
US-09-925-299-930
; Sequence 930, Application US/09925299
; Patent No. US2002005562A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 930
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (152)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (225)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-299-930

Query Match          69.4%; Score 25; DB 10; Length 308;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
|:|:|
Db 241 TDNLQAT 248

RESULT 13
US-09-765-272-122
; Sequence 122, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 520 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-09-765-272-122

Query Match          69.4%; Score 25; DB 10; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVL 5
|:|:|
Db 507 TNNVL 511

RESULT 14
US-09-965-313-7
; Sequence 7, Application US/09965313
; Patent No. US20020090680A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020090680A1 IL-9/TL-2 Receptor-Like Molecules
; FILE REFERENCE: 5800-17A
; CURRENT APPLICATION NUMBER: US/09/965,313
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 09/313,913
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-965-313-7

Query Match          69.4%; Score 25; DB 10; Length 522;
Best Local Similarity 66.7%; Pred. No. 4.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|:|:|
Db 58 TNNILR 63

RESULT 15
US-09-905-291A-250
; Sequence 250, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennle P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,291A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 250
; LENGTH: 546
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-905-291A-250

Query Match 69.4%; Score 25; DB 9; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6
Db 476 NNVLQ 480

Search completed: February 26, 2003, 15:42:46
Job time : 33 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:37:06 ; Search time 139 Seconds
(without alignments)
37.107 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US09_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	94.4	8	1	PCT-US98-13046-24
2	34	94.4	8	18	US-09-476-485A-24
3	34	94.4	8	24	US-10-045-353-24
4	34	94.4	123	18	US-09-476-485A-50
5	34	94.4	264	18	US-09-476-485A-2
6	34	94.4	270	1	PCT-US98-13046-2

7	34	94.4	270	24	US-10-045-353-2	Sequence 2, Appli
8	34	94.4	272	21	US-09-791-537-1173	Sequence 1173, Ap
9	34	94.4	279	18	US-09-476-485A-56	Sequence 56, Appl
10	34	94.4	279	21	US-09-791-537-109055	Sequence 109055,
11	34	94.4	286	1	PCT-US98-13046-23	Sequence 23, Appl
12	34	94.4	286	18	US-09-476-485A-23	Sequence 23, Appl
13	34	94.4	286	24	US-10-045-353-23	Sequence 23, Appl
14	34	94.4	303	18	US-09-476-485A-6	Sequence 6, Appli
15	31	86.1	73	19	US-09-513-996A-77260	Sequence 77260, A
16	31	86.1	74	19	US-09-513-996A-1504	Sequence 1504, Ap
17	31	86.1	79	19	US-09-513-996A-77259	Sequence 77259, A
18	31	86.1	80	19	US-09-513-996A-1503	Sequence 1503, Ap
19	31	86.1	90	19	US-09-513-996A-77258	Sequence 77258, A
20	31	86.1	91	19	US-09-513-996A-1502	Sequence 1502, Ap
21	31	86.1	104	12	US-08-827-356-4896	Sequence 4896, Ap
22	31	86.1	104	20	US-09-611-529-6497	Sequence 6497, Ap
23	30	83.3	25	21	US-09-724-429-79655	Sequence 79655, A
24	30	83.3	25	21	US-09-724-429-79656	Sequence 79656, A
25	30	83.3	25	21	US-09-724-429-80231	Sequence 80231, A
26	30	83.3	25	21	US-09-724-429-80232	Sequence 80232, A
27	30	83.3	25	21	US-09-724-429-81202	Sequence 81202, A
28	30	83.3	25	21	US-09-724-429-81203	Sequence 81203, A
29	30	83.3	25	21	US-09-724-429-81204	Sequence 81204, A
30	30	83.3	25	21	US-09-724-429-82066	Sequence 82066, A
31	30	83.3	25	21	US-09-724-429-82067	Sequence 82067, A
32	30	83.3	25	21	US-09-724-429-82068	Sequence 82068, A
33	30	83.3	25	21	US-09-724-429-83409	Sequence 83409, A
34	30	83.3	25	21	US-09-724-429-83410	Sequence 83410, A
35	30	83.3	25	21	US-09-724-429-83411	Sequence 83411, A
36	30	83.3	25	21	US-09-724-429-83412	Sequence 83412, A
37	30	83.3	25	21	US-09-724-429-84561	Sequence 84561, A
38	30	83.3	25	21	US-09-724-429-84562	Sequence 84562, A
39	30	83.3	25	21	US-09-724-429-84563	Sequence 84563, A
40	30	83.3	25	21	US-09-724-429-84564	Sequence 84564, A
41	30	83.3	25	21	US-09-724-429-86503	Sequence 86503, A
42	30	83.3	25	21	US-09-724-429-86504	Sequence 86504, A
43	30	83.3	25	21	US-09-724-429-86505	Sequence 86505, A
44	30	83.3	25	21	US-09-724-429-86506	Sequence 86506, A
45	30	83.3	25	21	US-09-724-429-86507	Sequence 86507, A

ALIGNMENTS

RESULT 1
PCT-US98-13046-24
; Sequence 24, Application PC/TUS9813046
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: USA
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/13046
; FILING DATE: June 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: June 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601

```

; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US98-13046-24

Query Match          94.4%; Score 34; DB 1; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 1 TNNVLOXT 8

RESULT 2
US-09-476-485A-24
; Sequence 24, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 24
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Dolichos lablab
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = any amino acid.
US-09-476-485A-24

Query Match          94.4%; Score 34; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 1 TNNVLOXT 8

RESULT 3
US-10-045-353-24
; Sequence 24, Application US/10045353
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/045,353
; FILING DATE: 29-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-045-353-24

Query Match          94.4%; Score 34; DB 24; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 1 TNNVLOXT 8

RESULT 4
US-09-476-485A-50
; Sequence 50, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Beta-subunit of D1-FRIL.
US-09-476-485A-50

Query Match          94.4%; Score 34; DB 18; Length 123;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
Db 27 TNNVLOXT 34

RESULT 5
US-09-476-485A-2
; Sequence 2, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.

```

; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
 ; TITLE OF INVENTION: and Products of Their Use
 ; FILE REFERENCE: 108236.119
 ; CURRENT APPLICATION NUMBER: US/09/476.485A
 ; CURRENT FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 08/881,189
 ; PRIOR FILING DATE: 1997-06-24
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 264
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: D1-FRIL.
 US-09-476-485A-2

Query Match 94.4%; Score 34; DB 18; Length 264;
 Best Local Similarity 87.5%; Pred. No. 42;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
 Db 27 TNNVLQXT 34
 ||||| |

RESULT 6
 PCT-US98-13046-2
 ; Sequence 2, Application PC/TUS9813046
 ; GENERAL INFORMATION:
 ; APPLICANT: Colucci et al.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 6900 Jericho Turnpike
 ; CITY: Syosset
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11791
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US98/13046
 ; FILING DATE: June 23, 1998
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/881,189
 ; FILING DATE: June 24, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 270 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-045-353-2
 ; Sequence 2, Application US/10045353
 ; GENERAL INFORMATION:
 ; APPLICANT: Colucci et al.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
 ; PROGENITOR CELL PRESERVATION FACTOR
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11753
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/045.353
 ; FILING DATE: 29-Oct-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/881,189
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 270 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 PCT-US98-13046-2

Query Match 94.4%; Score 34; DB 1; Length 270;
 Best Local Similarity 87.5%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
 Db 27 TNNVLQXT 34
 ||||| |

Db 27 TNNVLQXT 34
 RESULT 7
 US-10-045-353-2
 ; Sequence 2, Application US/10045353
 ; GENERAL INFORMATION:
 ; APPLICANT: Colucci et al.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
 ; PROGENITOR CELL PRESERVATION FACTOR
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11753
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/045.353
 ; FILING DATE: 29-Oct-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/881,189
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Feit, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 270 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 94.4%; Score 34; DB 24; Length 270;
 Best Local Similarity 87.5%; Pred. No. 43;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
 Db 27 TNNVLQXT 34
 ||||| |

RESULT 8
 US-09-791-537-1173
 ; Sequence 1173, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Blonmox, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1173
 ; LENGTH: 272
 ; TYPE: PRT
 ; ORGANISM: Dolichos lablab

US-09-791-537-1173

Query Match 94.4%; Score 34; DB 21; Length 272;
Best Local Similarity 87.5%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
| | | | | | |
Db 35 TNNVLQVT 42

RESULT 9

US-09-476-485A-56
; Sequence 56, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476.485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 56
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PVPRIIL.
US-09-476-485A-56

Query Match 94.4%; Score 34; DB 18; Length 279;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
| | | | | | |
Db 27 TNNVLQLT 34

RESULT 10

US-09-791-537-109055
; Sequence 109055, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 109055
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-09-791-537-109055

Query Match 94.4%; Score 34; DB 21; Length 279;
Best Local Similarity 87.5%; Pred. No. 44;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
| | | | | | |
Db 27 TNNVLQLT 34

RESULT 11

PCT-US98-13046-23
; Sequence 23, Application PC/TUS9813046
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: USA
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/13046
; FILING DATE: June 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: June 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
PCT-US98-13046-23

Query Match 94.4%; Score 34; DB 1; Length 286;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
| | | | | | |
Db 49 TNNVLQVT 56

RESULT 12

US-09-476-485A-23
; Sequence 23, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476.485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SpDLA.
US-09-476-485A-23

Query Match 94.4%; Score 34; DB 18; Length 286;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | | | |
Db 49 TNNVLOXT 56

RESULT 13
US-10-045-353-23
; Sequence 23, Application US/10045353
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/045,353
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/881,189
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: 381-44 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-10-045-353-23

Query Match 94.4%; Score 34; DB 24; Length 286;
Best Local Similarity 87.5%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | | | |
Db 49 TNNVLOXT 56

RESULT 14
US-09-476-485A-6
; Sequence 6, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; FILE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189

; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pv-FRIL.
US-09-476-485A-6

Query Match 94.4%; Score 34; DB 18; Length 303;
Best Local Similarity 87.5%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | | | |
Db 27 TNNVLOXT 34

RESULT 15
US-09-513-996A-77260
; Sequence 77260, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 77260
; LENGTH: 73
; TYPE: PRT

; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..73 /
; OTHER INFORMATION: any n or Xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..73 / Ceres Seq. ID 2104968
US-09-513-996A-77260

Query Match 86.1%; Score 31; DB 19; Length 73;
Best Local Similarity 75.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLOXT 8
| | | | | | |
Db 14 TNNVLOXT 21

Search completed: February 26, 2003, 15:41:37
Job time : 140 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:37:41 ; Search time 23 Seconds
(without alignments)
31.790 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	34	94.4	8	US-10-190-258A-11	Sequence 11, Appl
2	34	94.4	264	US-10-083-936B-2	Sequence 2, Appl
3	34	94.4	264	US-10-190-258A-2	Sequence 2, Appl
4	34	94.4	279	US-10-190-258A-6	Sequence 6, Appl
5	34	94.4	303	US-10-083-936B-6	Sequence 6, Appl
6	31	86.1	104	US-09-950-084-6497	Sequence 376, App
7	27	75.0	328	US-10-287-274-376	Sequence 6497, Ap
8	27	75.0	365	PCT-US02-35143-13	Sequence 13, Appl
9	27	75.0	365	US-10-286-326-13	Sequence 13, Appl
10	27	75.0	1912	PCT-US02-21361-48	Sequence 48, Appl
11	27	75.0	1912	US-10-188-186-48	Sequence 48, Appl
12	26	72.2	257	PCT-US02-36123-1120	Sequence 1120, Ap
13	26	72.2	262	US-09-950-084-4917	Sequence 4917, Ap
14	26	72.2	310	US-10-157-104-77	Sequence 77, Appl
15	26	72.2	310	US-10-005-041A-49	Sequence 49, Appl
16	26	72.2	310	US-10-005-041B-49	Sequence 49, Appl
17	26	72.2	313	US-10-157-104-73	Sequence 73, Appl
18	26	72.2	313	US-10-005-041A-50	Sequence 50, Appl
19	26	72.2	313	US-10-005-041A-108	Sequence 108, App
20	26	72.2	313	US-10-005-041B-50	Sequence 50, Appl
21	26	72.2	313	US-10-005-041B-108	Sequence 108, App
22	26	72.2	357	PCT-US02-36123-1122	Sequence 1122, Ap
23	26	72.2	357	US-10-157-104-75	Sequence 75, Appl
24	26	72.2	357	US-10-157-104-76	Sequence 76, Appl
25	26	72.2	357	US-10-005-041A-12	Sequence 12, Appl
26	26	72.2	357	US-10-005-041A-47	Sequence 47, Appl

27 26 72.2 357 6 US-10-005-041A-48 Sequence 48, Appl
28 26 72.2 357 6 US-10-005-041A-109 Sequence 109, App
29 26 72.2 357 6 US-10-005-041A-111 Sequence 111, App
30 26 72.2 357 6 US-10-005-041B-12 Sequence 12, Appl
31 26 72.2 357 6 US-10-005-041B-47 Sequence 47, Appl
32 26 72.2 357 6 US-10-005-041B-48 Sequence 48, Appl
33 26 72.2 357 6 US-10-005-041B-109 Sequence 109, App
34 26 72.2 357 6 US-10-005-041B-111 Sequence 111, App
35 26 72.2 400 1 PCT-US02-36122-22 Sequence 22, Appl
36 26 72.2 400 1 PCT-US02-36123-1124 Sequence 1124, Ap
37 26 72.2 418 1 PCT-US02-00667A-53 Sequence 53, Appl
38 26 72.2 631 6 US-10-310-154-436 Sequence 436, App
39 26 72.2 779 6 US-10-072-012-564 Sequence 564, App
40 26 72.2 779 6 US-10-072-012-731 Sequence 731, App
41 26 72.2 1136 6 US-10-072-012-280 Sequence 280, App
42 26 72.2 1136 6 US-10-072-012-729 Sequence 729, App
43 26 72.2 1165 6 US-10-072-012-730 Sequence 730, App
44 26 72.2 1178 6 US-10-310-154-551 Sequence 551, App
45 25 69.4 64 6 US-10-320-646-31 Sequence 31, Appl

ALIGNMENTS

RESULT 1
US-10-190-258A-11
; Sequence 11, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190, 258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 8
; TYPE: PPT
; ORGANISM: Dolichos lablab
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 7
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-190-258A-11

Query Match 94.4%; Score 34; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 1 TNNVLQXT 8

RESULT 2
US-10-083-936B-2
; Sequence 2, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/302,716
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2

```
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-083-936B-2

Query Match          94.4%; Score 34; DB 6; Length 264;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 3
US-10-190-258A-2
; Sequence 2, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-190-258A-2

Query Match          94.4%; Score 34; DB 6; Length 264;
Best Local Similarity 87.5%; Pred. No. 1.5;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 4
US-10-190-258A-6
; Sequence 6, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-190-258A-6

Query Match          94.4%; Score 34; DB 6; Length 279;
Best Local Similarity 87.5%; Pred. No. 1.6;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 5
US-10-083-936B-6
; Sequence 6, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/302,716
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-083-936B-6

Query Match          94.4%; Score 34; DB 6; Length 303;
Best Local Similarity 87.5%; Pred. No. 1.8;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy 1 TNNVLQXT 8
Db 27 TNNVLQVT 34

RESULT 6
US-09-950-084-6497
; Sequence 6497, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6497
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-6497

Query Match          86.1%; Score 31; DB 5; Length 104;
Best Local Similarity 75.0%; Pred. No. 2.6;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 1 TNNVLQXT 8
```


Db 71 TNNVLEIT 78

RESULT 7

US-10-287-274-376
; Sequence 376, Application US/10287274
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlson, Kari
; APPLICANT: Zyskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008DVI
; CURRENT APPLICATION NUMBER: US/10/287,274
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: US 09/711164
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 376
; LENGTH: 328
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-287-274-376

Query Match 75.0%; Score 27; DB 6; Length 328;
Best Local Similarity 66.7%; Pred. No. 94;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 83 TNNLIQ 88

RESULT 8

PCT-US02-35143-13
; Sequence 13, Application PC/TUS0235143
; GENERAL INFORMATION:
; APPLICANT: San, Ka-Yui
; APPLICANT: Berrios-Rivera, Susana
; APPLICANT: Bennett, George
; TITLE OF INVENTION: Recycling System for Manipulation of Intracellular NADH Availability
; FILE REFERENCE: P02328WO
; CURRENT APPLICATION NUMBER: PCT/US02/35143
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/335,371
; PRIOR FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Rhodococcus
PCT-US02-35143-13

Query Match 75.0%; Score 27; DB 1; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 338 TNNLLQ 343

RESULT 9

US-10-286-326-13
; Sequence 13, Application US/10286326
; GENERAL INFORMATION:
; APPLICANT: San, Ka-Yui
; APPLICANT: Berrios-Rivera, Susana
; APPLICANT: Bennett, George

; TITLE OF INVENTION: Recycling System for Manipulation of Intracellular NADH Availability
; FILE REFERENCE: P02328US1
; CURRENT APPLICATION NUMBER: US/10/286,326
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/335,371
; PRIOR FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Rhodococcus
US-10-286-326-13

Query Match 75.0%; Score 27; DB 6; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 338 TNNLLQ 343

RESULT 10

PCT-US02-21361-48
; Sequence 48, Application PC/TUS0221361
; GENERAL INFORMATION:
; APPLICANT: CuraGen Corporation, et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: PCT/US02/21361
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 48
; LENGTH: 1912
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-21361-48

Query Match 75.0%; Score 27; DB 1; Length 1912;
Best Local Similarity 71.4%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8
|||:|
Db 1355 NNVIQRT 1361

RESULT 11

US-10-188-186-48
; Sequence 48, Application US/10188186

GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 48
; LENGTH: 1912
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-186-48

Query Match 75.0%; Score 27; DB 6; Length 1912;
Best Local Similarity 71.4%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLOXT 8
Db 1355 NNVQRT 1361

RESULT 12
PCT-US02-36123-1120
; Sequence 1120, Application PC/TUS0236123
; GENERAL INFORMATION:
; APPLICANT: American Cyanamid Company, and Fletcher, Leah D., McMichael, John C.,
; TITLE OF INVENTION: Russell, David P., and Zagursky, Robert J.
; FILE REFERENCE: Antigen, Immunogenic Compositions and Uses Thereof
; CURRENT APPLICATION NUMBER: PCT/US02/36123
; CURRENT FILING DATE: 2003-01-02
; NUMBER OF SEQ ID NOS: 6653
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1120
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Alloiococcus otitidis
PCT-US02-36123-1120

Query Match 72.2%; Score 26; DB 1; Length 257;
Best Local Similarity 71.4%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLOXT 8
Db 166 NNVQRT 172

RESULT 13

US-09-950-084-4917
; Sequence 4917, Application US/09950084
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963US2
; CURRENT APPLICATION NUMBER: US/09/950,084
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 4917
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-950-084-4917

Query Match 72.2%; Score 26; DB 5; Length 262;
Best Local Similarity 62.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 64 TNTVIQRT 71

RESULT 14
US-10-157-104-77
; Sequence 77, Application US/10157104
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Burgess, Catherine E
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R.
; APPLICANT: Stone, David J.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-694 CIP2CON2
; CURRENT APPLICATION NUMBER: US/10/157,104
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/185,674
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,535

;; PRIOR FILING DATE: 2000-02-28
;; PRIOR APPLICATION NUMBER: 60/186,585
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,604
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,584
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,717
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,716
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,719
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/186,827
;; PRIOR FILING DATE: 2000-03-03
;; PRIOR APPLICATION NUMBER: 60/218,323
;; PRIOR FILING DATE: 2000-07-14
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 104
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 77
;; LENGTH: 310
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-157-104-77

Query Match 72.2%; Score 26; DB 6; Length 310;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 154 SNSVLOST 161

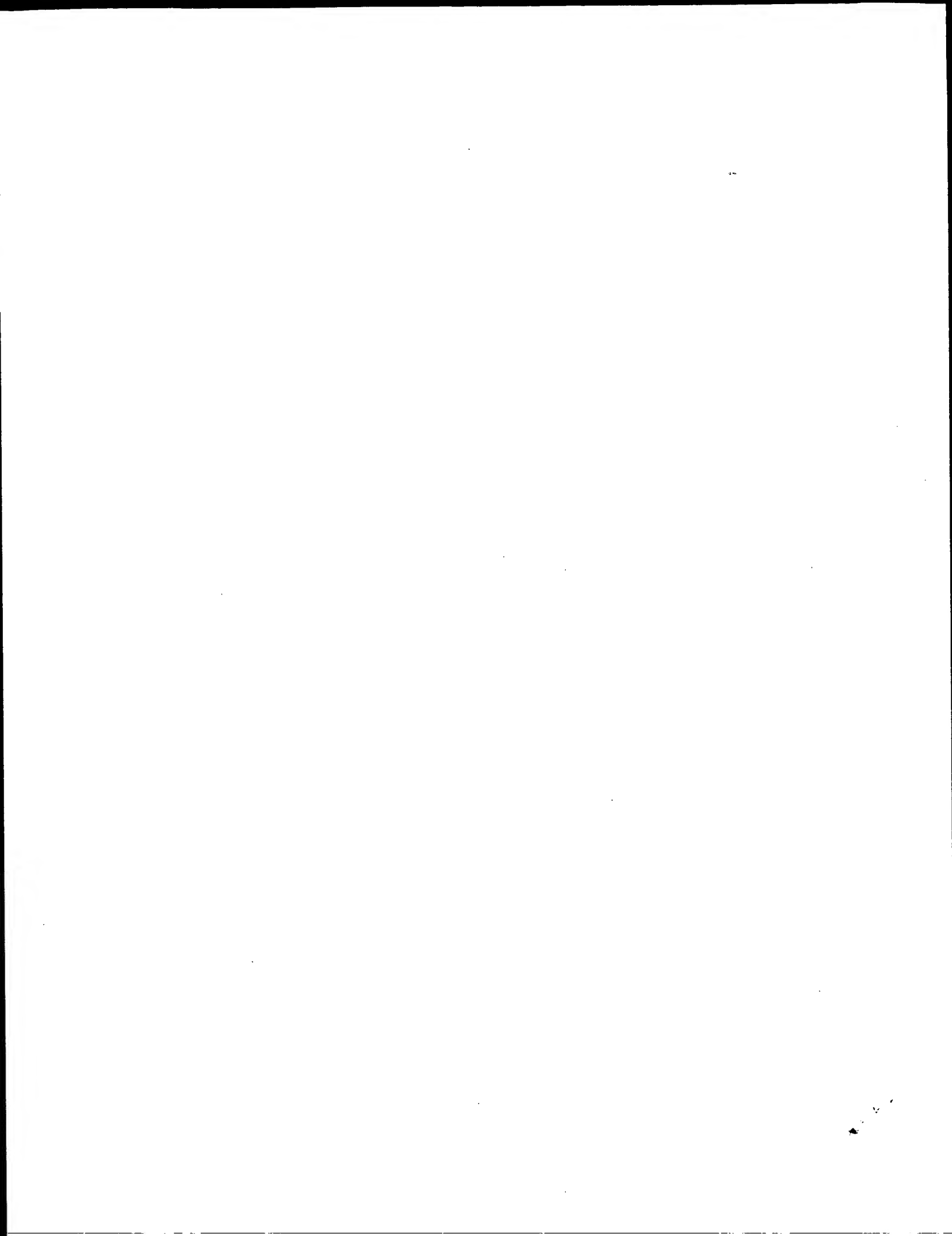
RESULT 15
US-10-005-041A-49
Sequence 49. Application US/10005041A
GENERAL INFORMATION:
APPLICANT: Casman, Stacie J
APPLICANT: Padigar, Muralidhara
APPLICANT: Burgess, Catherine E
APPLICANT: Shinkets, Richard A
APPLICANT: Spytek, Kimberly A
APPLICANT: Gilbert, Jennifer A
APPLICANT: Mayotte, Jane E
APPLICANT: Baumgartner, Jason C
APPLICANT: Mishra, Vishnu
APPLICANT: Vernet, Corine AM
APPLICANT: Dickinson, Kevin S
APPLICANT: Ballinger, Robert A
APPLICANT: Wolenc, Adam R
APPLICANT: Edinger, Shlomit R
APPLICANT: Macbougall, John R
APPLICANT: Smithson, Glennnda
APPLICANT: Ellerman, Karen
APPLICANT: Stone, David J
APPLICANT: Gunther, Erik
APPLICANT: Gerlach, Valerie
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-215
CURRENT APPLICATION NUMBER: US/10/005,041A
CURRENT FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 60/251,459
PRIOR FILING DATE: 2000-12-05
PRIOR APPLICATION NUMBER: 60/259,007
PRIOR FILING DATE: 2000-12-29
NUMBER OF SEQ ID NOS: 205
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 49
LENGTH: 310
TYPE: PRT
ORGANISM: Homo sapiens

US-10-005-041A-49

Query Match 72.2%; Score 26; DB 6; Length 310;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
Db 154 SNSVLOST 161

Search completed: February 26, 2003, 15:42:06
Job time : 23 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:36:26 ; Search time 30 seconds
(without alignments)
7.846 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents, AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length	ID	Description
1	34	94.4	8	4	US-08-881-189B-24
2	34	94.4	270	4	US-08-881-189B-2
3	34	94.4	286	4	US-08-881-189B-23
4	29	80.6	1184	4	US-09-541-782-2
5	29	80.6	1184	4	US-09-723-820-2
6	27	75.0	365	1	US-08-089-755A-5
7	27	75.0	365	1	US-08-421-754-5
8	27	75.0	365	2	US-08-421-791-5
9	27	75.0	365	2	US-08-851-088-4
10	27	75.0	365	4	US-08-851-089-10
11	26	72.2	903	4	US-09-228-986-78
12	25	69.4	20	3	US-08-872-094-12
13	25	69.4	188	4	US-08-737-109-14
14	25	69.4	239	2	US-08-665-647-11
15	25	69.4	261	4	US-09-045-973-3
16	25	69.4	299	4	US-09-134-001C-3433
17	25	69.4	307	1	US-08-164-614A-11
18	25	69.4	307	2	US-08-456-489B-11
19	25	69.4	464	2	US-09-021-323-3
20	25	69.4	480	3	US-09-189-035-5
21	25	69.4	480	4	US-09-382-086-5
22	25	69.4	520	4	US-08-961-083-122
23	25	69.4	522	1	US-08-164-614A-10
24	25	69.4	522	2	US-08-456-489B-10
25	25	69.4	536	1	US-08-164-614A-12
26	25	69.4	536	2	US-08-456-489B-12
27	25	69.4	580	1	US-08-309-512-6

28 25 69.4 580 5 PCT-US92-08756A-6 Sequence 6, Appli
29 25 69.4 599 1 US-08-222-619-2 Sequence 2, Appli
30 25 69.4 599 4 US-08-221-767-24 Sequence 24, Appli
31 25 69.4 599 5 PCT-US95-04075-2 Sequence 2, Appli
32 25 69.4 622 1 US-08-547-197-1 Sequence 1, Appli
33 25 69.4 622 4 US-08-957-940-1 Sequence 11, Appli
34 25 69.4 630 4 US-09-228-986-71 Sequence 11, Appli
35 25 69.4 693 1 US-08-463-620-11 Sequence 11, Appli
36 25 69.4 693 2 US-08-224-917-11 Sequence 11, Appli
37 25 69.4 693 2 PCT-US95-03934A-11 Sequence 11, Appli
38 25 69.4 693 5 US-08-368-459A-8 Sequence 8, Appli
39 25 69.4 921 2 US-08-487-826B-8 Sequence 8, Appli
40 25 69.4 921 4 US-09-210-288-8 Sequence 8, Appli
41 25 69.4 997 3 US-08-872-094-8 Sequence 8, Appli
42 25 69.4 1040 4 US-08-961-083-118 Sequence 118, Appli
43 25 69.4 1481 2 US-08-616-844-40 Sequence 40, Appli
44 25 69.4 1481 2 US-08-599-654-40 Sequence 40, Appli
45

ALIGNMENTS

RESULT 1
US-08-881-189B-24
; Sequence 24, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-881-189B-24

Query Match 94.4%; Score 34; DB 4; Length 8;
Best Local Similarity 87.5%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8

Db 1 TNNVLQXT 8

```

RESULT 2
US-08-881-189B-2
; Sequence 2, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-2

Query Match 94.4%; Score 34; DB 4; Length 270;
Best Local Similarity 87.5%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
DB 27 TNNVLQVT 34

RESULT 3
US-08-881-189B-23
; Sequence 23, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS

```

```

; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-881-189B-23

```

```

Query Match 94.4%; Score 34; DB 4; Length 286;
Best Local Similarity 87.5%; Pred. No. 4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TNNVLQXT 8
DB 49 TNNVLQVT 56

```

```

RESULT 4
US-09-541-782-2
; Sequence 2, Application US/09541782
; Patent No. 6284480
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/541.782
; CURRENT FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1184
; TYPE: PRT
; ORGANISM: A. nidulans
US-09-541-782-2

```

```

Query Match 80.6%; Score 29; DB 4; Length 1184;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TNNVLQXT 8
DB 521 TNDVLQVT 528

```

```

RESULT 5
US-09-723-820-2
; Sequence 2, Application US/09723820
; Patent No. 6468760
; GENERAL INFORMATION:
; APPLICANT: Nislow, Corey
; APPLICANT: Sakowicz, Roman
; APPLICANT: Beraud, Christophe
; TITLE OF INVENTION: Antifungal Assay
; FILE REFERENCE: 1015
; CURRENT APPLICATION NUMBER: US/09/723,820
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/541,782

```

;; PRIOR FILING DATE: 2000-04-03
;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 2
;; LENGTH: 1184
;; TYPE: PRT
;; ORGANISM: A. nidulans
US-09-723-820-2

Query Match 80.6%; Score 29; DB 4; Length 1184;
Best Local Similarity 75.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
|||
DB 521 TNDVLQQT 528

RESULT 6
US-08-089-755A-5
; Sequence 5, Application US/08089755A
; Patent No. 5356801
; GENERAL INFORMATION:
; APPLICANT: Rambossek, John
; APPLICANT: Piddington, Chris
; APPLICANT: Kovacevich, Brian R
; APPLICANT: Young, Kevin D
; APPLICANT: Denome, Sylvia A
; TITLE OF INVENTION: Recombinant DNA Encoding A
; TITLE OF INVENTION: Desulfurization Biocatalyst
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/089,755A
FILING DATE: 09-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/911,845
FILING DATE: 10-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: EBC92-03A
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-089-755A-5

Query Match 75.0%; Score 27; DB 1; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||
DB 338 TNNLLQ 343

RESULT 7
US-08-421-754-5
; Sequence 5, Application US/08421754
; Patent No. 5578478
; GENERAL INFORMATION:
; APPLICANT: Rambossek, John
; APPLICANT: Piddington, Chris
; APPLICANT: Kovacevich, Brian R
; APPLICANT: Young, Kevin D
; APPLICANT: Denome, Sylvia A
; TITLE OF INVENTION: Recombinant DNA Encoding A
; TITLE OF INVENTION: Desulfurization Biocatalyst
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,754
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/304,081
FILING DATE: 01-SEP-1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/089,755
FILING DATE: 09-JUL-1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/911,845
FILING DATE: 10-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: EBC92-03AZZ
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-754-5

Query Match 75.0%; Score 27; DB 1; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||
DB 338 TNNLLQ 343

RESULT 8
US-08-421-791-5
; Sequence 5, Application US/08421791
; Patent No. 5879914
; GENERAL INFORMATION:
; APPLICANT: Rambossek, John
; APPLICANT: Piddington, Chris
; APPLICANT: Kovacevich, Brian R

APPLICANT: Young, Kevin D
APPLICANT: Denome, Sylvia A
TITLE OF INVENTION: Recombinant DNA Encoding A
TITLE OF INVENTION: Desulfurization Biocatalyst
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith and Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/421,791
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/304,081
FILING DATE: 01-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/089,755
FILING DATE: 09-JUL-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/911,845
FILING DATE: 10-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: EBC92-03A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-9540
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-421-791-5

Query Match 75.0%; Score 27; DB 2; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 338 TNNLLQ 343

RESULT 9
US-08-851-088-4
Sequence 4, Application US/08851088
Patent No. 5952208
GENERAL INFORMATION:
APPLICANT: Darzins, Aldis
APPLICANT: XI, Lei
APPLICANT: Childs, John D.
APPLICANT: Monticello, Daniel J.
APPLICANT: Squires, Charles H.
TITLE OF INVENTION: DS2 Gene Expression In Pseudomonas Hosts
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts

COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,088
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/835,185
FILING DATE: 07-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Elmore, Carolyn S.
REGISTRATION NUMBER: 37,567
REFERENCE/DOCKET NUMBER: EBC96-06A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-851-088-4

Query Match 75.0%; Score 27; DB 2; Length 365;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 338 TNNLLQ 343

RESULT 10
US-08-851-089-10
Sequence 10, Application US/08851089
Patent No. 6133016
GENERAL INFORMATION:
APPLICANT: Aldis Darzins
APPLICANT: Gregory T. Mrachko
TITLE OF INVENTION: A Spingomonas Biodesulfurization
Catalyst
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/851,089
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/835,292
FILING DATE: 07-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Elmore, Carolyn S.
REGISTRATION NUMBER: 37,567
REFERENCE/DOCKET NUMBER: EBC97-06A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240

TELEFAX: (781) 861-9540
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 365 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-851-089-10

Query Match 75.0%; Score 27; DB 4; Length 365;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 ||:||
 Db 338 TNNLLQ 343

RESULT 11

US-09-228-986-78
 Sequence 78, Application US/09228986
 Patent No. 6359198

GENERAL INFORMATION:

APPLICANT: Strabala, Timothy

APPLICANT: Nieuwenhuizen, Niels

TITLE OF INVENTION: Compositions Isolated from Plant Cells

TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling

FILE REFERENCE: 11000/1020

CURRENT APPLICATION NUMBER: US/09/228,986

CURRENT FILING DATE: 1999-01-12

NUMBER OF SEQ ID NOS: 130

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 78

LENGTH: 903

TYPE: PRT

ORGANISM: Eucalyptus grandis

US-09-228-986-78

Query Match 72.2%; Score 26; DB 4; Length 903;
 Best Local Similarity 71.4%; Pred. No. 7.9e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 NNVLQXT 8
 ||:||
 Db 492 NNLLQGT 498

RESULT 12

US-08-872-094-12
 Sequence 12, Application US/08872094
 Patent No. 6020128

GENERAL INFORMATION:

APPLICANT: Steiner, Bret M.

TITLE OF INVENTION: DNA Polymerase from Treponema pallidum

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Jones & Askew

STREET: 191 Peachtree Street, 37th Floor

CITY: Atlanta

STATE: Georgia

COUNTRY: U.S.A.

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/872,094

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Warren, William L.

REGISTRATION NUMBER: 36,714
 REFERENCE/DOCKET NUMBER: 03063-0320
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404) 818-3700
 TELEFAX: (404) 818-3799
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-872-094-12

Query Match 69.4%; Score 25; DB 3; Length 20;
 Best Local Similarity 62.5%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 TNNVLQXT 8
 ||:||
 Db 8 TNDTLQMT 15

RESULT 13

US-08-737-109-14

Sequence 14, Application US/08737109

Patent No. 6455688

GENERAL INFORMATION:

APPLICANT: SLABAS, Antoni Ryszard

APPLICANT: ELBOROUGH, Kieran Michael

APPLICANT: BRIGHT, Simon William Jonathan

APPLICANT: FENTEN, Philip Anthony

TITLE OF INVENTION: Plant Gene Specifying Acetyl Coenzyme A

TITLE OF INVENTION: Carboxylase and Transformed Plants Containing Same

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pillsbury Madison & Sutro, L.L.P.

STREET: 1100 New York Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MS Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/737,109

FILING DATE: 21-OCT-1996

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB94/00846

FILING DATE: 02-MAY-1994

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 188 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

FRAGMENT TYPE: N-terminal

ORGANISM: Arabidopsis thaliana

IMMEDIATE SOURCE:

CLONE: pKL081

US-08-737-109-14

Query Match 69.4%; Score 25; DB 4; Length 188;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVL 5
Db 26 TNNVL 30

RESULT 14

US-08-665-647-11
; Sequence 11, Application US/08665647
; Patent No. 5935803
; GENERAL INFORMATION:
; APPLICANT: Dasquez, Nicki J.
; APPLICANT: Ron, Dorit
; APPLICANT: Voronova, Anna F.
; APPLICANT: Napolitano, Eugene W.
; TITLE OF INVENTION: METHODS TO IDENTIFY IMMUNOMODULATORS
; TITLE OF INVENTION: USING COGNATE INTERACTION OF PKC-THETA
; NUMBER OF SEQUENCES: 89
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW - Ste. 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,647
; FILING DATE: 18-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22550-20025.25
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 822-0168
; TELEX: 90-4030 MRSNFOERSWSH
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-647-11

Query Match 69.4%; Score 25; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NNVLQ 6
Db 26 NNVLQ 30

RESULT 15

US-09-045-973-3
; Sequence 3, Application US/09045973
; Patent No. 6165767
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Yue, Henry
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah
; TITLE OF INVENTION: PROTEIN PHOSPHATASE RELATED MOLECULES
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive

; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,973
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0491 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT18
; CLONE: 2534680
US-09-045-973-3

Query Match 69.4%; Score 25; DB 4; Length 261;
Best Local Similarity 62.5%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
Db 204 TNNILFT 211

Search completed: February 26, 2003, 15:39:11
Job time : 31 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:40 ; Search time 16 Seconds
(without alignments)
84.118 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	83.3	105	2 B54864	mannose/glucose-sp
2	43	65.2	249	2 JQ1981	lectin II - Scotch
3	42	63.6	408	2 F63379	conserved hypothet
4	39	59.1	290	2 S66355	lectin-related sto
5	39	59.1	924	2 C83572	organic solvent to
6	38	57.6	467	2 T16319	hypothetical prote
7	38	57.6	705	2 S23352	gene unc-93 protei
8	38	57.6	708	2 T19369	hypothetical prote
9	38	57.6	807	2 T19836	hypothetical prote
10	37	56.1	24	2 S08292	lectin - smooth ra
11	37	56.1	220	2 T39548	hypothetical prote
12	37	56.1	298	2 S77536	lactose transport
13	37	56.1	560	2 S54093	hypothetical prote
14	36	54.5	284	2 T23636	hypothetical prote
15	36	54.5	425	2 C87619	2-isopropylmalate
16	36	54.5	524	2 B70479	hypothetical prote
17	36	54.5	534	2 T30629	hypothetical prote
18	36	54.5	590	2 F75501	hypothetical prote
19	36	54.5	699	2 T33375	sensor histidine k
20	35	53.0	124	2 S67606	hypothetical prote
21	35	53.0	190	2 A56666	probable membrane
22	35	53.0	191	2 JH0605	neurocalcin - frui
23	35	53.0	191	2 A48979	neural visinin-lik
24	35	53.0	193	2 JH0616	visinin-like prote
25	35	53.0	193	2 JH0816	neurocalcin (clone
26	35	53.0	193	2 I50676	neural visinin-lik
27	35	53.0	193	2 S47565	gene Rem-1 protein
28	35	53.0	203	2 T25609	calcium-binding pr
29	35	53.0	266	2 S51833	hypothetical prote
					arcelin-4 precursor

ALIGNMENTS

RESULT 1

B54864

mannose/glucose-specific lectin beta chain - Dolichos lab lab (field bean)

C:Species: Dolichos lab lab (field bean)

C:Date: 23-Jun-1995 #sequence_revision 10-Nov-1995 #text_change 07-May-1999

C:Accession: B54864; B48591

R:Gowda, L.R.; Savithri, H.S.; Rao, D.R.

J. Biol. Chem. 269, 18789-18793, 1994

A:Title: The complete primary structure of a unique mannose/glucose-specific lectin f

A:Reference number: A54864; MUID:94308133; PMID:8034631

A:Accession: B54864

A:Molecule type: protein

A:Residues: 1-105 <GOW>

C:Keywords: glycoprotein; lectin

Query Match

Best Local Similarity 83.3%; Score 55; DB 2; Length 105;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFD 12

Db 1 AQSLSFXFTKFD 12

RESULT 2

JQ1981

lectin II - Scotch broom

N:Alternate names: 2-acetamido-2-deoxy-D-galactose binding lectin II

C:Species: Cytisus scoparius (Scotch broom)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999

C:Accession: JQ1981

R:Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T.

J. Biochem. 112, 366-375, 1992

A:Title: The primary structure of the Cytisus scoparius seed lectin and a carbohydrate

A:Reference number: JQ1981; MUID:93054441; PMID:1429525

A:Accession: JQ1981

A:Molecule type: protein

A:Residues: 1-249 <KON>

A:Experimental source: seed

C:Superfamily: plant lectin

C:Keywords: lectin

Query Match

Best Local Similarity 65.2%; Score 43; DB 2; Length 249;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14

Db 1 SEELSFSTKFTD 14

RESULT 3

F83379
conserved hypothetical protein PA2127 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
R:Accession: F83379
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83379
A:Molecule type: DNA
A:Status: preliminary
A:Cross-references: GB:AE004640; GB:AE004091; NID:g9948139; PIDN:AAG05515.1; GSPDB:GN0011
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA2127
C:Superfamily: Escherichia coli ybdN protein

Query Match 63.6%; Score 42; DB 2; Length 408;
Best Local Similarity 57.1%; Pred. No. 3.3;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 385 ARSLSFQFSQADID 398
|:|||||:|:|
1 AQSLSFXFTKFDLD 14

RESULT 4
S66355
lectin-related storage protein precursor - Cladragis lutea (fragment)
C:Species: Cladragis lutea
C:Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 20-Aug-1999
C:Accession: S66355; S66299
R:van Damme, E.J.M.; Barre, A.; Bemer, V.; Rouge, P.; van Leuven, F.; Peumans, W.J.
Plant Mol. Biol. 29, 579-598, 1995
A:Title: A lectin and a lectin-related protein are the two most prominent proteins in the seed of *Cladragis lutea*
A:Reference number: S66299; MUID:96123235; PMID:8534854
A:Accession: S66355
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-290 <VAN>
A:Cross-references: EMBL:U21940; NID:g1141754; PIDN:AAC49150.1; PID:g1141755
A:Accession: S66299
A:Molecule type: protein
A:Residues: 37-56 <VAM>
A:Experimental source: bark
C:Comment: This lectin-related protein has no carbohydrate binding activity.
C:Superfamily: plant lectin
C:Keywords: lectin
F:1-36/Domain: signal sequence (fragment) #status predicted <SIG>
F:37-290/Product: lectin-related storage protein #status experimental <MAT>

Query Match 59.1%; Score 39; DB 2; Length 290;
Best Local Similarity 70.0%; Pred. No. 8.4;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 38 EALSFTFTKF 47
|:|||||
2 QSLSPXFTKFK 11

RESULT 5
CB3572
organic solvent tolerance protein OstA precursor PA0595 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: CB3572
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: CB3572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-924 <STO>
A:Cross-references: GB:AE004495; GB:AE004091; NID:g9946459; PIDN:AAG03984.1; GSPDB:GN0011
A:Experimental source: strain PA01
C:Genetics:
A:Gene: ostA; PA0595

Query Match 59.1%; Score 39; DB 2; Length 924;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Db 569 TLKLYTKYDLD 580
|:|:|:|:|
3 SLSFXFTKFDLD 14

RESULT 6
TL6319
hypothetical protein F41C3.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: TL6319
R:Chissoe, S.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid F41C3.
A:Reference number: Z18494
A:Accession: TL6319
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-467 <CHI>
A:Cross-references: EMBL:U23521; NID:g746536; PID:g746538; PIDN:AAC46809.1; CESP:F41C3
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F41C3.2
A:Introns: 40/1; 415/2

Query Match 57.6%; Score 38; DB 2; Length 467;
Best Local Similarity 58.3%; Pred. No. 22;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Db 52 SLNFKFEQD 63
|:|:|:|:|
3 SLSFXFTKFDLD 14

RESULT 7
S23352
gene unc-93 protein 1 - Caenorhabditis elegans
N:Contains: gene unc-93 protein 2
C:Species: Caenorhabditis elegans
C:Date: 03-Mar-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C:Accession: S23352; S23353
R:Levin, J.Z.; Horvitz, H.R.
J. Cell Biol. 117, 143-155, 1992
A:Title: The Caenorhabditis elegans unc-93 gene encodes a putative transmembrane protein
A:Reference number: S23352; MUID:92210636; PMID:1313436
A:Accession: S23352
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-705 <LEV>
A:Cross-references: EMBL:X64415; NID:g6909; PIDN:CAA45760.1; PID:g6910
A:Accession: S23353
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 6-705 <LE2>
A:Cross-references: EMBL:X64415; NID:g6909; PIDN:CAA45761.1; PID:g6911
C:Genetics:
A:Introns: 36/3; 118/2; 147/2; 215/2; 247/3; 284/2; 329/2; 376/2; 418/2; 449/1; 478/2

F:6-705/Product: gene unc-93 protein 2 #status predicted <MAT2>

Query Match 57.6%; Score 38; DB 2; Length 705;
Best Local Similarity 70.0%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKFK 11
||| | |||

Db 623 QSLQFAFTKY 632

RESULT 8

T19969
hypothetical protein C46F11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19969

R:Burton, J.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19206

A:Accession: T19969

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-708 <WIL>

A:Cross-references: EMBL:Z81449; PIDN:CA803760.1; GSPDB:GN000021; CESP:C46F11.1

A:Experimental source: clone C46F11

C:Genetics:

A:Gene: CESP:C46F11.1

A:Map position: 3

A:Introns: 36/3; 118/2; 147/2; 215/2; 247/3; 284/2; 329/2; 369/1; 421/2; 452/1; 481/2; 5

Query Match

Best Local Similarity 57.6%; Score 38; DB 2; Length 708;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKFK 11

||| | |||

Db 626 QSLQFAFTKY 635

RESULT 9

T19836

hypothetical protein C38H2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T19836

R:Slims, M.

submitted to the EMBL Data Library, August 1994

A:Reference number: Z19185

A:Accession: T19836

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-807 <WIL>

A:Cross-references: EMBL:Z35641; PIDN:CAA84706.1; GSPDB:GN000021; CESP:C38H2.1

A:Experimental source: clone C38H2

C:Genetics:

A:Gene: CESP:C38H2.1

A:Map position: 3

A:Introns: 42/2; 81/3; 190/2; 215/1; 257/3; 270/1; 388/1; 443/2; 461/2; 528/3; 571/3; 61

Query Match

Best Local Similarity 57.6%; Score 38; DB 2; Length 807;

Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQLSFXTKFDLD 14

||| | | |

Db 704 AQLSFAFRWFLD 717

RESULT 10

S08292

lectin - smooth rattlebox (fragment)

C:Species: Crotalaria pallida (smooth rattlebox)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1993

C:Accession: S08292

R:Nguyen, O.K.; Guillaume, J.L.; Hoebeke, J.

Biochim. Biophys. Acta 1033, 210-213, 1990

A:Title: A blood group A specific lectin from the seeds of *Crotalaria striata*.

A:Reference number: S08292; MUID:90167102; PMID:2306467

A:Accession: S08292

A:Molecule type: protein

A:Residues: 1-24 <NGO>

A:Note: the source is designated as *Crotalaria striata*

C:Keywords: lectin

Query Match 56.1%; Score 37; DB 2; Length 24;

Best Local Similarity 70.0%; Pred. No. 1.5;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXTKFDLD 14

|| | | |

Db 5 SFSTKPFSTD 14

RESULT 11

T39548

hypothetical protein SPBC16A3.04 - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T39548

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A:Reference number: Z21862

A:Accession: T39548

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-220 <WOO>

A:Cross-references: EMBL:AL021748; PIDN:CAAL6855.1; GSPDB:GN000067; SPDB:SPBC16A3.04

A:Experimental source: strain 972h-; cosmid c16A3

C:Genetics:

A:Gene: SPDB:SPBC16A3.04

A:Map position: 2

Query Match

Best Local Similarity 56.1%; Score 37; DB 2; Length 220;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQLSFXTKFDLD 14

|| | | |

Db 167 AQLGAVTKSDLE 180

RESULT 12

S77536

lactose transport system permease protein lacF - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein slr1202

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S77536

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77536

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-298 <KAN>

A:Cross-references: EMBL:D90905; GB:AB001339; NID:g1652360; PIDN:BAA17383.1; PID:g165

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: lacF

C:Superfamily: inner membrane protein ugpA

C:Keywords: binding protein-dependent transport system; lactose transport; membrane p

Query Match 56.1%; Score 37; DB 2; Length 298;
Best Local Similarity 46.2%; Pred. No. 21;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
II: I II::II
Db 37 QAFSLSFYQYELD 49

RESULT 13
S54093
hypothetical protein YPR072w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YP9499.27
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 19-Apr-2002
C:Accession: S54093; S69060
R:Badcock, K.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54059
A:Accession: S54093
A:Molecule type: DNA
A:Residues: 1-560 <BAD>
A:Cross-references: EMBL:Z49219; NID:g805025; PIDN:CAA89189.1; PID:g805052; MIPS:YPR072w
A:Experimental source: Strain AB972
R:Couch, J.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of S. cerevisiae cosmid 9513.
A:Reference number: S69057
A:Accession: S69060
A:Molecule type: DNA
A:Residues: 1-560 <COU>
A:Cross-references: EMBL:U51033; NID:g1230676; PIDN:AAB69123.1; PID:g1230680; MIPS:YPR072w
C:Genetics:
A:Gene: SGD:NOT5
A:Cross-references: SGD:S0006276
A:Map position: 16R
C:Superfamily: hypothetical protein YPR072w

Query Match 56.1%; Score 37; DB 2; Length 560;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14
II: I II::II
Db 464 SLARIFMKFDLD 475

RESULT 14
T23636
hypothetical protein M01B2.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 26-May-2000
C:Accession: T23636
R:Lloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19776
A:Accession: T23636
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-284 <WIL>
A:Cross-references: EMBL:Z83116; PIDN:CAB05561.1; GSPDB:GN00023; CBSP:M01B2.3
A:Experimental source: clone M01B2
C:Genetics:
A:Gene: CBSP:M01B2.3
A:Map position: 5
A:Introns: 211/3
C:Superfamily: Caenorhabditis elegans hypothetical protein B0250.6

Query Match 54.5%; Score 36; DB 2; Length 284;
Best Local Similarity 46.2%; Pred. No. 31;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
II: I II::II
Db 40 QSMFIYWKFSVD 52

RESULT 15
C87619
hypothetical protein CC2989 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: C87619
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko, N.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <STO>
A:Cross-references: GB:AE005673; NID:gl3424623; PIDN:AAK24951.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2989

Query Match 54.5%; Score 36; DB 2; Length 425;
Best Local Similarity 63.6%; Pred. No. 47;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AOSLSFXFTKF 11
II: I II::II
Db 361 AQALSFVFTPY 371

Search completed: February 26, 2003, 14:49:22
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:40 ; Search time 11 Seconds
(without alignments)
52.788 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDL14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	83.3	237	1	LECA_DOLLA
2	43	65.2	248	1	LEC2_CYTSC
3	39	59.1	290	1	LEC2_CIALU
4	37	56.1	24	1	LEC_CROST
5	37	56.1	560	1	NOT5_YEAST
6	36	54.5	520	1	LEU1_AQUAE
7	35	53.0	189	1	NCAB_DRONE
8	35	53.0	190	1	VIS1_HUMAN
9	35	53.0	192	1	NCAD_CHICK
10	35	53.0	192	1	NCAD_HUMAN
11	35	53.0	192	1	NCAD_MOUSE
12	35	53.0	192	1	VIS3_CHICK
13	35	53.0	192	1	VIS3_HUMAN
14	35	53.0	192	1	VIS3_MOUSE
15	35	53.0	236	1	LECA_CRAFL
16	35	53.0	240	1	LECS_VATWA
17	35	53.0	266	1	ARC4_PHAVU
18	35	53.0	431	1	EF1G_DRONE
19	35	53.0	847	1	YAF6_YEAST
20	35	53.0	930	1	YMP8_CHLPN
21	34	51.5	165	1	NCAB_BOVIN
22	34	51.5	174	1	CALB_NEUCR
23	34	51.5	185	1	LEC_VICVI
24	34	51.5	190	1	VIS2_RAT
25	34	51.5	192	1	HIPP_HUMAN
26	34	51.5	192	1	NECX_APLCA
27	34	51.5	196	1	CAGS_HELPY
28	34	51.5	199	1	CAGS_HELPJ
29	34	51.5	243	1	LEC2_CYTSE
30	34	51.5	270	1	LEC2_SOPJA
31	34	51.5	285	1	LCB1_BOBPS
32	34	51.5	285	1	LCS1_BOBPS
33	34	51.5	285	1	LCS2_BOBPS

34	34	51.5	286	1	LCB2_BOBPS
35	34	51.5	292	1	LECS_SOPJA
36	34	51.5	363	1	YRUS_CAEEL
37	34	51.5	435	1	YCW9_YEAST
38	34	51.5	449	1	PCOL_HUMAN
39	34	51.5	575	1	11OR_MOUSE
40	34	51.5	706	1	MR11_MOUSE
41	34	51.5	706	1	MR11_RAT
42	34	51.5	708	1	MR11_HUMAN
43	33	50.0	15	1	LEC2_PSOSC
44	33	50.0	24	1	LEC_CROJUF
45	33	50.0	167	1	YMH6_CAEEL

ALIGNMENTS

RESULT 1
LECA_DOLLA
ID AC P38662; STANDARD; PRT: 237 AA.
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Lectin.
OS Dolichos lab lab (Field bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=35936;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Lignosus; TISSUE=Seed;
RX MEDLINE=94308133; PubMed=8034631;
RA Gowda L.R., Savithri H.S., Rajagopal Rao D.;
RT "The complete primary structure of a unique mannose/glucose-specific
lectin from field bean (Dolichos lab lab)";
RL J. Biol. Chem. 269:18789-18793(1994).
CC -1- FUNCTION: D-MANNOSE/D-GLUCOSE-BINDING LECTIN. REQUIRES CA2+ AND
CC MN2+ IONS FOR FULL ACTIVITY.
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains.
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR HSP; P02866; IONA.
DR InterPro: IPR000985; Lectin_legA.
DR InterPro: IPR001220; Lectin_legB.
DR Pfam: PF00138; lectin_legA; 1.
DR Pfam: PF00139; lectin_legB; 1.
DR ProDom: PD000671; Lectin_legA; 1.
DR ProDom: PD000711; Lectin_legB; 1.
DR PROSITE: PS00307; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE: PS00308; LECTIN_LEGUME_BETA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Acetylation.
FT CHAIN 1 105
FT CARBOHYD 106 237
FT MOD_RES 69 69
FT N-LINKED (GLCNAC...) (POTENTIAL).
FT METAL 115 115
FT METAL 117 117
FT METAL 120 120
FT METAL 122 122
FT METAL 127 127
FT METAL 132 132
SQ SEQUENCE 237 AA; 25718 MW; BEB7E84DC2895327 CRC64;

Query Match 83.3%; Score 55; DB 1; Length 237;
Best Local Similarity 91.7%; Pred. No. 0.0025;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
DB 1 AQSLSFXFTKFD 12

```

RESULT 2
LEC2_CYTSC
ID LEC2_CYTSC STANDARD; PRT; 248 AA.
AC P29257;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 2-acetamido-2-deoxy-D-galactose-binding seed lectin II (CSII).
OS Cytisus scoparius (Scotch broom).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Cytisus.
OX NCBI_TaxID=3835;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=93054441; PubMed=1429525;
RA Konami Y., Yamamoto K., Osawa T., Irimura T.;
RT "The primary structure of the Cytisus scoparius seed lectin and a
RT carbohydrate-binding peptide.";
RL J. Biochem. 112:366-375(1992).
RN [2]
RP SEQUENCE OF 1-32.
RX MEDLINE=84307428; PubMed=6477513;
RA Young N.M., Watson D.C., Williams R.E.;
RT "Structural differences between two lectins from Cytisus scoparius,
RT both specific for D-galactose and N-acetyl-D-galactosamine.";
RL Biochem. J. 222:41-48(1984).
CC -1- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR HSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
CC -1- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
DR EMBL; U01940; AAC49150.1; -.
DR HSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 36 LECTIN-RELATED PROTEIN.
FT CHAIN 37 290 MANGANESE (BY SIMILARITY).
FT METAL 161 161 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 163 163 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 167 167 CALCIUM (BY SIMILARITY).
FT METAL 170 170 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 175 175 MANGANESE (BY SIMILARITY).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 248 AA; 27028 MW; C960F457C8C9F62A CRC64;

Query Match 65.2%; Score 43; DB 1; Length 248;
Best Local Similarity 57.1%; Pred. No. 0.5;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
Db 1 SEELSFSTFKFTD 14

RESULT 3
LECR_CLALU
ID LECR_CLALU STANDARD; PRT; 290 AA.
AC Q39527;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Lectin-related protein precursor (CLLRP) (LRPCL) (Fragment).
OS Cladrastis lutea (Yellow wood).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Cladrastis.
OX NCBI_TaxID=38412;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 37-56.
RC TISSUE=Bark;

```

```

RX MEDLINE=96123235; PubMed=8534854;
RA van Damme E.J.M., Barre A., Rouge P., van Leuven F.,
RA Peumans W.J.;
RT "A lectin and a lectin-related protein are the two most prominent
RT proteins in the bark of yellow wood (Cladrastis lutea).";
RL Plant Mol. Biol. 29:579-598(1995).
CC -1- FUNCTION: DOES NOT HAVE ANY CARBOHYDRATE BINDING OR AGGLUTINATION
CC ACTIVITY.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: TO OTHER LEGUMINOUS LECTINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; U01940; AAC49150.1; -.
DR HSP; P19588; 1LUL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
KW Lectin; Calcium; Manganese; Glycoprotein; Signal.
FT NON_TER 1
FT SIGNAL <1 36 LECTIN-RELATED PROTEIN.
FT CHAIN 37 290 MANGANESE (BY SIMILARITY).
FT METAL 161 161 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 163 163 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 167 167 CALCIUM (BY SIMILARITY).
FT METAL 170 170 MANGANESE AND CALCIUM (BY SIMILARITY).
FT METAL 175 175 MANGANESE (BY SIMILARITY).
FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 290 AA; 31378 MW; 64F2DBE7B2E20B14 CRC64;

Query Match 59.1%; Score 39; DB 1; Length 290;
Best Local Similarity 70.0%; Pred. No. 3.3;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 QLSFXFTKFK 11
Db 38 EALSFTTKF 47

RESULT 4
LEC_CROST
ID LEC_CROST STANDARD; PRT; 24 AA.
AC P16351;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-MAY-1992 (Rel. 22, Last annotation update)
DE Lectin (Fragment).
OS Crotalaria striata (Smooth rattlebox) (Crotalaria pallida).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Crotalariaeae;
OC Crotalaria.
OX NCBI_TaxID=3830;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=90167102; PubMed=2306467;
RA Khang N.O., Guillaume J.-L., Hoebeke J.;
RT "A blood group A specific lectin from the seeds of Crotalaria
RT striata.";

```


RL Biochim. Biophys. Acta 1033:210-213(1990).
 CC -!- FUNCTION: AGGLUTINATES ERYTHROCYTES OF BLOOD GROUP A. BINDS IN
 CC DECREASING ORDER OF AFFINITY: N-ACETYL-D-GALACTOSAMINE,
 CC D-GALACTOSE, AND D-GALACTOSAMINE.
 CC
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC
 CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
 CC PIR: S08292; S08292.
 DR InterPro: IPR000985; Lectin_legA.
 DR InterPro: IPR001220; Lectin_legB.
 DR Pfam: PF00139; lectin_legB; 1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; PARTIAL.
 DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; PARTIAL.
 KW Lectin; Glycoprotein.
 FT NON_TER 24
 SQ SEQUENCE 24 AA; 2875 MW; DC62B82CD9F9BB66 CRC64;
 Query Match 56.1%; Score 37; DB 1; Length 24;
 Best Local Similarity 70.0%; Pred. No. 0.58;
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 5 SPXFTKFDLD 14
 Db 5 SFSFTKFDST 14
 RESULT 5
 NOT5_YEAST
 ID NOT5_YEAST STANDARD; PRT; 560 AA.
 AC Q12514;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE General negative regulator of transcription subunit 5.
 GN NOT5 OR YPR072W OR YP9499.27.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 FX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Cough J.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoerge W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Borstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Deilus H., DiPaolo T., Dubois E., Duesterhoeft A.,
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kaiman S., Kleine K.,
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,
 RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";
 RN Nature 387:103-105(1997).
 RX [3]
 RP CHARACTERIZATION.
 RX MEDLINE=98172735; PubMed=9511744;
 RA Oberholzer U., Collart M.A.;
 RT "Characterization of NOT5 that encodes a new component of the Not
 RT protein complex.";
 RL Gene 207:61-69(1998).
 CC -!- FUNCTION: NEGATIVELY REGULATES THE BASAL AND ACTIVATED

CC TRANSCRIPTION OF MANY GENES. PREFERENTIALLY AFFECTS TC-TYPE TATA
 CC ELEMENT-DEPENDENT TRANSCRIPTION. COULD DIRECTLY OR INDIRECTLY
 CC INHIBITS COMPONENT(S) OF THE GENERAL TRANSCRIPTION MACHINERY.
 CC -!- SUBUNIT: FORMS A COMPLEX THAT COMPRISE NOT1, NOT2, NOT3, NOT4 AND
 CC NOT5.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: TO YEAST NOT3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U51033; AAB68123.1; -;
 CC EMBL: Z71255; CAA94980.1; -;
 CC EMBL: Z49219; CAA89189.1; -;
 CC TRANSFAC: T03515; -;
 CC SGD: S0006276; YPR072W.
 KW Nuclear protein; Transcription regulation; Repressor; Coiled coil.
 FT DOMAIN 3 26 COILED COIL (POTENTIAL).
 FT DOMAIN 37 71 COILED COIL (POTENTIAL).
 FT DOMAIN 124 177 COILED COIL (POTENTIAL).
 SQ SEQUENCE 560 AA; 65854 MW; 07313711983E92BC CRC64;
 Query Match 56.1%; Score 37; DB 1; Length 560;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 Qy 3 SLSPXFTKFDLD 14
 Db 464 SLRIFMKFDLD 475
 RESULT 6
 LEU1_AQUAE
 ID LEU1_AQUAE STANDARD; PRT; 520 AA.
 AC O67862;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 2-isopropylmalate synthase (EC 4.1.3.12) (Alpha-isopropylmalate
 DE synthase) (Alpha-IPM synthetase).
 GN LEUA OR AQ_2090
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 CC Aquifex.
 CX [1]
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., AuJay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus".
 RL Nature 392:353-358(1998).
 CC -!- FUNCTION: Catalyzes the condensation of the acetyl group of
 CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
 CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
 CC -!- CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + H(2)O.
 CC acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O.
 CC -!- PATHWAY: Leucine biosynthesis; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE
 CC SYNTHASE FAMILY. LEUA 1 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AF000772; AAC07824.1; ALT_INIT.
DR InterPro; IPR002034; AIPM/Hcgt_synth.
DR InterPro; IPR000891; HMGL-like.
DR Pfam; PF00682; HMGL-like; 1.
DR TIGRFAMs; TIGR00973; leuA_bact; 1.
DR PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
DR PROSITE; PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
KW Leucine biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 520 AA; 58103 MW; 95B210FA872871D4 CRC64;

Query Match 54.5%; Score 36; DB 1; Length 520;
Best Local Similarity 53.8%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14
DB 343 EELGFKFTKEELD 355

RESULT 7

ID NCAH_DROME STANDARD; PRT; 189 AA.
AC P42325; Q9Y67;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurocalcin homolog.
GN NCA OR CG7641.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Canton-S;
RX MEDLINE=95081147; PubMed=7989365;
RZ Teng D.H.-F., Chen C.-K., Hurley J.B.;
RT "A highly conserved homologue of bovine neurocalcin in Drosophila
RT melanogaster is a Ca(2+)-binding protein expressed in neuronal
RT tissues";
RL J. Biol. Chem. 269:31900-31907(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Fendell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [3]
RP MYRISTOYLATION, AND FUNCTION.
RX MEDLINE=96215323; PubMed=8626592;
RA Faurobert E., Chen C.-K., Hurley J.B., Teng D.H.-F.;
RT "Drosophila neurocalcin, a fatty acylated, Ca2+-binding protein that
RT associates with membranes and inhibits in vitro phosphorylation of
RT bovine rhodopsin";
RL J. Biol. Chem. 271:10256-10262(1996).
CC -!- FUNCTION: INHIBITS THE PHOSPHORYLATION OF RHODOPSIN IN A CALCIUM-
CC DEPENDENT MANNER. PROBABLY BINDS TWO OR THREE CALCIUM IONS.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN NEURONAL TISSUES. HIGH LEVEL
CC EXPRESSION SEEN IN THE CORTICAL REGIONS OF THE CENTRAL BRAIN AND
CC LOWER LEVELS IN THE LAMINA, THE FIRST OPTIC LOBE OF THE BRAIN. IT
CC IS ALSO FOUND IN THE THORACIC GANGLIA.
CC -!- DEVELOPMENTAL STAGE: FOUND IN THE EMBRYOS, LARVAE AND PUPAE.
CC EXPRESSION IN THE ADULT HEADS IS HIGHER THAN IN THE BODIES.
CC -!- MASS SPECTROMETRY: MW=21975.51; METHOD=Electrospray.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL; U15735; AAA62152.1; -.
DR EMBL; AE003515; AAF49082.1; -.

DR HSSP; P36610; 1G81.

DR FlyBase; FBgn0013303; Nca.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR001125; Recoverin.

DR Pfam; PF00036; efhand; 3.

DR PRINTS; PR00450; RECOVERIN.

DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM000054; EFh; 3.

DR PROSITE; PS00018; EF_HAND; 3.

KW Calcium-binding; Repeat; Myristate.

FT INIT_MET 0 0

FT LIPID 1 1 MYRISTATE.

FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1.

FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).

FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).

FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).

FT SEQUENCE 189 AA; 21762 MW; 3715201BEA2824F CRC64;

Query Match 53.0%; Score 35; DB 1; Length 189;

Best Local Similarity 46.2%; Pred. No. 12;

Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14

DB 98 OKLKWAFSMYDLD 110

```

RESULT 8
VISI_HUMAN
ID VISI_HUMAN STANDARD; PRT; 190 AA.
AC 28677; P29103; Q9UM20; P42323;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Visinin-like protein 1 (VILIP-1) (Neural visinin-like protein 1) (NVL-
DE 1) (NVP-1) (21 kDa CABP) (Neurocalcin alpha) (Hippocalcin-like protein
DE 3) (HLP3).
GN VSNL1 OR VISL1.
OS Homo sapiens (Human).
OS Mus musculus (Mouse).
OS Rattus norvegicus (Rat).
OS Bos taurus (Bovine).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 10090, 10116, 9913, 9031;
RN
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=96079121; PubMed=8530085;
RA Polymopoulos M.H., Ide S., Soares M.B., Lennon G.G.;
RT "Sequence characterization and genetic mapping of the human VSNL1
RT gene, a homologue of the rat visinin-like peptide RNVp1.";
RL Genomics 29:273-275(1995).
RN
RP
RC SEQUENCE FROM N.A.
RC SPECIES=Human;
RA Bellingham J.;
RT "Peptide conservation between avian and mammalian visinin-like
RT proteins.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC SPECIES=Human;
RX MEDLINE=99449057; PubMed=10520747;
RA Kobayashi M., Sakai E., Furuta Y., Takamatsu K.;
RT "Isolation of two human cDNAs, HLP3 and HLP4, homologous to the
RT neuron-specific calcium-binding protein genes.";
RL DNA Seq. 9:171-176(1998).
RN
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC SPECIES=Rat; TISSUE=Brain;
RX MEDLINE=92272711; PubMed=1375457;
RA Kuno T., Kajimoto Y., Hashimoto T., Mukai H., Shirai Y., Saheki S.,
RA Tanaka C.;
RT "cDNA cloning of a neural visinin-like Ca(2+)-binding protein.";
RL Biochem. Biophys. Res. Commun. 184:1219-1225(1992).
RN
RP SEQUENCE FROM N.A.
RC SPECIES=Mouse; STRAIN=BALB/C;
RA Kuno T.;
RT Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Brain;
RX MEDLINE=98228226; PubMed=9560316;
RA Kato M., Watanabe Y., Iino S., Takaoka Y., Kobayashi S., Haga T.,
RA Hidaka H.;
RT "Cloning and expression of a cDNA encoding a new neurocalcin isoform
RT (neurocalcin alpha) from bovine brain.";
RL Biochem. J. 331:871-876(1998).
RN
RP SEQUENCE FROM N.A.
RC SPECIES=Chicken; STRAIN=white leghorn; TISSUE=Brain;
RX MEDLINE=93061804; PubMed=1359372;
RA Lenz S., Henschel Y., Zopf D., Voss B., Gundelfinger E.D.;
RT "VILIP, a cognate protein of the retinal calcium binding proteins
RT visinin and recoverin, is expressed in the developing chicken
RT brain.";
RL Brain Res. Mol. Brain Res. 15:133-140(1992).

```

```

[8]
RN SEQUENCE OF 7-26 AND 99-129.
RP SPECIES=Bovine; TISSUE=Brain;
RX MEDLINE=92406915; PubMed=1527077;
RA Terasawa M., Nakano A., Kobayashi R., Hidaka H.;
RT "Neurocalcin: a novel calcium-binding protein from bovine brain.";
RL J. Biol. Chem. 267:19596-19599(1992).
CC -1- FUNCTION: REGULATES (IN VITRO) THE INHIBITION OF RHODOPSIN
CC PHOSPHORYLATION IN A CALCIUM-DEPENDENT MANNER.
CC -1- TISSUE SPECIFICITY: BRAIN AND RETINA. NEURON-SPECIFIC IN THE
CC CENTRAL AND PERIPHERAL NERVOUS SYSTEM.
CC -1- MISCELLANEOUS: PROBABLY BINDS THREE CALCIUM IONS.
CC -1- SIMILARITY: TO OTHER EF-HAND THREE CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14747; AAA91295.1; -
CC EMBL; AF039555; AAD02174.1; -
CC EMBL; AB001104; BAA86891.1; -
CC EMBL; D10666; BAA01517.1; -
CC EMBL; D21165; BAA04701.1; -
CC EMBL; AB006006; BAA28716.1; -
CC EMBL; X63530; CAA45093.1; -
CC PIR; JH0605; JH0605.
CC PIR; S18904; S18904.
CC PIR; A48979; A48979.
CC HSSP; P36610; ICB1.
CC
CC Genew; HGNC:12722; VSNL1.
CC MIM; 600817; -
CC MGI; 1349453; Vsnll.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR001125; Recoverin.
CC Pfam; PF00036; efhand; 6.
CC PRINTS; PR00450; RECOVERIN.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00054; EFh; 3.
CC PROSITE; PS00018; EF_HAND; 3.
CC
CC KW Calcium-binding; Repeat; Myristate.
FT INIT_MET 0
FT LIPID 1
FT DOMAIN 35 46
FT CA_BIND 72 83
FT CA_BIND 108 119
FT CA_BIND 158 169
FT CONFLICT 24 24
FT CONFLICT 64 64
FT CONFLICT 74 74
FT CONFLICT 117 117
FT CONFLICT 139 139
FT CONFLICT 157 157
FT CONFLICT 168 168
FT CONFLICT 171 171
SQ SEQUENCE 190 AA; 22011 MW; 954A3BE018A149D8 CRC64;
BY SIMILARITY.
MYRISTATE (BY SIMILARITY).
ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
EF-HAND 2 (POTENTIAL).
EF-HAND 3 (POTENTIAL).
EF-HAND 4 (POTENTIAL).
H -> A (IN REF. 8).
A -> G (IN REF. 1).
N -> I (IN REF. 1).
R -> P (IN REF. 8).
M -> K (IN REF. 5).
M -> K (IN REF. 3).
D -> G (IN REF. 1).
K -> R (IN REF. 1).
Query Match 53.0%; Score 35; DB 1; Length 190;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 QSLSFXTKFDLD 14
| | | |
Db 98 QKLWAFNMYDLD 110
RESULT 9
NCAD_CHICK STANDARD; PRT; 192 AA.
ID NCAD_CHICK

```

```

AC 012953;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurocalcin delta.
GN NCALD.
OS Gallus gallus (Chicken), and
OS Poephila guttata (zebra finch) (Taeniopygia guttata).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031, 59729;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Chicken; TISSUE=Gizzard;
RA Schonekess B.O., Walsh M.P.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=P.guttata; TISSUE=Telencephalon;
RA Wade J., Peabody C.T., Smith G.W.;
RT "Sexually dimorphic pattern of neurocalcin expression in the
RT developing zebra finch telencephalon."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS (BY
CC similarity).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U91630; AAB51149.1; -.
DR EMBL; AF272896; AAG09045.1; -.
DR HSSP; P36610; IG81.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001125; Recoverin.
DR Pfam; PF00036; ehand; 3
DR PRINTS; PR00450; RECOVERIN.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00054; EFh; 3.
DR PROSITE; PS00018; EF_HAND; 3.
KW Calcium-binding; Repeat; Myristate; Multigene family.
FT INIT_MET 0
FT BY SIMILARITY.
FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 192 AA; 22130 MW; 807254352EC57003 CRC64;
Query Match 53.0%; Score 35; DB 1; Length 192;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 2 QLSLFXFTKFDLD 14
DB 98 QKLKWAFSMYDLD 110
| | | | |
| | | | |
RESULT 10
NCAD HUMAN STANDARD; PRT; 192 AA.
ID NCAD HUMAN
AC P29554; Q9H0W2;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Neurocalcin delta.
GN NCALD.
OS Homo sapiens (Human), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606, 9913;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Fetal brain;
RA MEDLINE=21167385; PubMed=11267673;
RA Wang W., Zhou Z., Zhao W., Huang Y., Tang R., Ying K., Xie Y., Mao Y.;
RT "Molecular cloning, mapping and characterization of the human
RT neurocalcin delta gene (NCALD).";
RL Biochim. Biophys. Acta 1518:162-167(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Human; TISSUE=Brain;
RA MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mexes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [3]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 106-111 AND 120-125.
RC SPECIES=Bovine; TISSUE=Brain;
RA MEDLINE=92287085; PubMed=1599450;
RA Okazaki K., Watanabe M., Ando Y., Hagiwara M., Terasawa M., Hidaka H.;
RT "Full sequence of neurocalcin, a novel calcium-binding protein
RT abundant in central nervous system.";
RL Biochem. Biophys. Res. Commun. 185:147-153(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=Bovine; TISSUE=Brain;
RA MEDLINE=93247712; PubMed=8387172;
RA Hidaka H., Okazaki K.;
RT "Neurocalcin family: a novel calcium-binding protein abundant in
RT bovine central nervous system.";
RL Neurosci. Res. 16:73-77(1993).
RN [5]
RP CALCIUM-BINDING DATA.
RC SPECIES=Bovine;
RA MEDLINE=95153407; PubMed=7852401;
RA Ladant D.;
RT "Calcium and membrane binding properties of bovine neurocalcin delta
RT expressed in Escherichia coli.";
RL J. Biol. Chem. 270:3179-3185(1995).
RN [6]
RP CRYSTALLIZATION.
RC SPECIES=Bovine;
RA MEDLINE=96407688; PubMed=8811741;
RA Kumar V.D., Hidaka H., Okazaki K., Vijay-Kumar S.;
RT "Crystallization and preliminary X-ray crystallographic studies of
RT recombinant bovine neurocalcin delta.";
RL Proteins 25:261-264(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).
RC SPECIES=Bovine;
RA MEDLINE=99101387; PubMed=9886296;
RA Vijay-Kumar S., Kumar V.D.;
RT "Crystal structure of recombinant bovine neurocalcin.";
RL Nat. Struct. Biol. 6:80-88(1999).
CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS.
CC -!- TISSUE SPECIFICITY: RETINA, CEREBRUM, CEREBELLUM, BRAIN STEM,
CC SPINAL CORD, TESTIS, OVARY AND SMALL INTESTINE.
CC -!- MISCELLANEOUS: FIVE ISOPROTEIN FORMS OF NEUROCALCIN ARE DESIGNATED
CC ALPHA, BETA, GAMMA1, GAMMA2, AND DELTA.
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO

```

```

CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF251061; AA34951.1; -
CC EMBL: AL136612; CAB66547.1; -
CC EMBL: D10884; BAA01706.1; -
CC PIR: JH0616; JH0616.
CC PDB: 1BJF; 22-JUL-99.
CC Genew: HGNC:7655; NCALD.
CC MIM: 606722; -
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 3.
CC PRINTS: PR00450; RECOVERIN.
CC ProDom: PD000012; EF-hand; 1.
CC SMART: SM00054; EFh; 3.
CC PROSITE: PS00018; EF_HAND; 3.
CC KW Calcium-binding; Repeat; Myristate; 3D-structure; Multigene family.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE.
CC FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
CC FT CA_BIND 72 83 EF_HAND 2 (POTENTIAL).
CC FT CA_BIND 108 119 EF_HAND 3 (POTENTIAL).
CC FT CA_BIND 156 167 EF_HAND 4 (POTENTIAL).
CC FT CONFLICT 122 122 E -> V (IN REF. 2).
CC SQ SEQUENCE 192 AA; 22114 MW; 807242F9F4D3BCD9 CRC64;

Query Match 53.0%; Score 35; DB 1; Length 192;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
DB 98 QKLKWFMSMYDLD 110

RESULT 11
NCAD_MOUSE STANDARD; PRT; 192 AA.
ID NCAD_MOUSE
AC Q91X97;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Neurocalcin delta.
GN NCALD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Salivary gland;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
CC RHODOPSIN PHOSPHORYLATION. BINDS THREE CALCIUM IONS (BY
CC similarity).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: BC011162; AAH11162.1; -
CC InterPro: IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 3.
CC ProDom: PD000012; EF-hand; 1.
CC PROSITE: PS00018; EF_HAND; 3.
CC KW Calcium-binding; Repeat; Myristate; Multigene family.
CC FT INIT_MET 0 0 BY SIMILARITY.
CC FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
CC FT CA_BIND 72 83 EF_HAND 2 (POTENTIAL).
CC FT CA_BIND 108 119 EF_HAND 3 (POTENTIAL).
CC FT CA_BIND 156 167 EF_HAND 4 (POTENTIAL).
CC SQ SEQUENCE 192 AA; 22100 MW; 807246EDA592BCD9 CRC64;

Query Match 53.0%; Score 35; DB 1; Length 192;
Best Local Similarity 46.2%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
DB 98 QKLKWFMSMYDLD 110

RESULT 12
VIS3_CHICK STANDARD; PRT; 192 AA.
ID VIS3_CHICK
AC P42324;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Rem-1 protein (Hippocalcin-like protein 1).
GN HPCAL1 OR REM1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=95215060; PubMed=7700627;
RA Kraut N., Frampton J., Graf T.;
RT "Rem-1, a putative direct target gene of the Myb-Ets fusion
RT oncoprotein in haematopoietic progenitors, is a member of the
RT recoverin family.";
RL Oncogene 10:1027-1036(1995).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN HEMOPOIETIC CELLS, BUT ALSO IN
CC THE FIBROBLASTS, BONE MARROW, BRAIN, EYE AND GUT.
CC -!- INDUCTION: BY THE MYB-ETS FUSION ONCOGENE.
CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY
CC SIMILARITY).
CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
CC THE RECOVERIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X80875; CAA56843.1; -
CC HSPSP; P36610; IG81.
CC InterPro: IPR002048; EF-hand.
CC InterPro: IPR001125; Recoverin.
CC Pfam: PF00036; ehand; 3.
CC PRINTS: PR00450; RECOVERIN.
CC ProDom: PD000012; EF-hand; 1.
CC SMART: SM00054; EFh; 3.
CC PROSITE: PS00018; EF_HAND; 3.

```

KW Calcium-binding; Repeat; Myristate.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (BY SIMILARITY).
 FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
 FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
 SQ SEQUENCE 192 AA; BE7221EA537332FA CRC64;

Query Match 53.0%; Score 35; DB 1; Length 192;
 Best Local Similarity 46.2%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14
 | | | | : | | |
 Db 98 QKLKWAFFSMYDLD 110

RESULT 13
 VIS3_HUMAN STANDARD; PRT; 192 AA.
 AC P37235; Q969S5;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Visinin-like protein 3 (VILIP-3) (Calcium-binding protein BDR-1)
 DE (HLP2) (Hippocalcin-like protein 1).
 GN HPCALL1 OR BDR1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94312451; PubMed=8038222;
 RA Kobayashi M., Takamatsu K., Fujishiro M., Saitoh S., Noguchi T.;
 RT "Molecular cloning of a novel calcium-binding protein structurally
 related to hippocalcin from human brain and chromosomal mapping of
 its gene.";
 RL Biochim. Biophys. Acta 1222:515-518(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain, and Pancreas;
 RA Strausberg R.;
 RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
 RHODOPSIN PHOSPHORYLATION.
 CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY
 SIMILARITY).
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
 THE RECOVERIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D16227; BAA03754.1; -
 CC DR EMBL; BC009846; AAH09846.1; -
 CC DR EMBL; BC017028; AAH17028.1; -
 CC DR EMBL; BC017482; AAH17482.1; -
 CC DR HSP; P36610; IG81.
 CC DR Genew; HGNC:5145; HPCALL1.
 CC DR MIM; 600207; -
 CC DR InterPro; IPR002048; EF-hand.
 CC DR InterPro; IPR001125; Recoverin.
 CC DR Pfam; PF00036; efhand; 3.
 CC DR PRINTS; PR00450; RECOVERIN.
 CC DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM00054; EFh; 3.
 DR PROSITE; PS00018; EF_HAND; 3.
 KW Calcium-binding; Repeat; Myristate.
 FT INIT_MET 0 0 BY SIMILARITY.
 FT LIPID 1 1 MYRISTATE (POTENTIAL).
 FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1.
 FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
 FT CONFLICT 18 18 N -> K (IN REF. 1).
 FT CONFLICT 87 87 A -> G (IN REF. 1).
 FT CONFLICT 92 93 SR -> RG (IN REF. 1).
 SQ SEQUENCE 192 AA; 277320AEDC02AE79 CRC64;

Query Match 53.0%; Score 35; DB 1; Length 192;
 Best Local Similarity 46.2%; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSPXFTKFDLD 14
 | | | | : | | |
 Db 98 QKLKWAFFSMYDLD 110

RESULT 14
 VIS3_MOUSE STANDARD; PRT; 192 AA.
 AC P35333;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Visinin-like protein 3 (VILIP-3) (Neural visinin-like protein 3)
 DE (NVL-3) (NVP-3) (Hippocalcin-like protein 1).
 GN HPCALL1.
 OS Mus musculus (Mouse), and
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; STRAIN=BALE/c; TISSUE=Brain;
 RA Ritter B., Modregger J., Plomann M.;
 RT "Interactions of the murine neural visinin-like protein 3 (mNVP-3).";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse; TISSUE=Breast tumor;
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Rat; TISSUE=Brain;
 RX MEDLINE=93367470; PubMed=8360675;
 RA Kajimoto Y., Shirai Y., Mukai H., Kuno T., Tanaka C.;
 RT "Molecular cloning of two additional members of the neural visinin-
 RT like Ca(2+)-binding protein gene family.";
 RL J. Neurochem. 61:1091-1096(1993).
 CC -!- FUNCTION: MAY BE INVOLVED IN THE CALCIUM-DEPENDENT REGULATION OF
 RHODOPSIN PHOSPHORYLATION.
 CC -!- TISSUE SPECIFICITY: IN NEURONAL CELLS, BUT NOT AS SPECIFICALLY AS
 VILIP-1 OR VILIP-2.
 CC -!- MISCELLANEOUS: PROBABLY BINDS TWO OR THREE CALCIUM IONS (BY
 SIMILARITY).
 CC -!- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BELONGS TO
 THE RECOVERIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

```

CC EMBL; AF085192; AAC35552.1; -
DR EMBL; BC001997; AA001997.1; -
DR EMBL; D13126; BAA02428.1; -
DR PIR; JH0816; JH0816.
DR HSP; P36610; IG81.
DR MGD; MGI:1855689; Hpcall.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF000138; LECTIN_LEG.
DR PRINTS; PR00450; RECOVERIN.
DR PRODOM; PD000012; EF-hand; 1.
DR SMART; SM00054; EFH; 3.
DR PROSITE; PS00018; EF-HAND; 3.
KW Calcium-binding; Repeat; Myristate.
FT INIT_MET 0 BY SIMILARITY.
FT LIPID 1 MYRISTATE (BY SIMILARITY).
FT DOMAIN 35 46 ANCESTRAL CALCIUM SITE 1 (POTENTIAL).
FT CA_BIND 72 83 EF-HAND 2 (POTENTIAL).
FT CA_BIND 108 119 EF-HAND 3 (POTENTIAL).
FT CA_BIND 156 167 EF-HAND 4 (POTENTIAL).
SQ SEQUENCE 192 AA; 22207 MW; 178ADDD3F89B097B CRC64;

```

Query Match 53.08; Score 35; DB 1; Length 192;
 Best Local Similarity 46.24; Pred. No. 12;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 2 QSLSFXTKFDLD 14
   | | | | |
DB 98 OKLWAFSMYDLD 110

```

```

RESULT 15
LECA_CRAFL
ID LECA_CRAFL STANDARD; PRT; 236 AA.
AC P81517; P81636;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE LECTIN ALPHA CHAIN [Contains: LECTIN BETA CHAIN; LECTIN GAMMA CHAIN].
OS Cratylia floribunda.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Cratylia.
OX NCBI_TaxID=83131;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RA Cavada B.S., Nogueira N.A.P., Farias C.M.A.S., Grangeiro T.B.,
RA Romas M.V., Thole H.H., Raida M., Rouge P., Calvete J.J.;
RT "Primary structure and kinetic interaction with glycoproteins of the
RT lectin from seeds of Cratylia floribunda."
RL Protein Pept. Lett. 6:27-34(1999).
RN [2]
RP SEQUENCE, MASS SPECTROMETRY, AND X-RAY CRYSTALLOGRAPHY (3.5
RP ANGSTROMS).
RC TISSUE=Seed;
RX MEDLINE=99184792; PubMed=10082964;
RA Calvete J.J., Thole H.H., Raida M., Urbanke C., Romero A.,
RA Grangeiro T.B., Ramos M.V., Almeida da Rocha I.M., Guimaraes F.N.,
RA Cavada B.S.;
RT "Molecular characterization and crystallization of Diocleinae
RT lectins."
RL Biochim. Biophys. Acta 1430:367-375(1999).
CC -!- FUNCTION: D-Mannose/D-Glucose-binding lectin. Mixture of 60% alpha
CC lectin and 40% of its beta and gamma proteolytic fragments.
CC -!- COFACTOR: Requires calcium and manganese ions for full activity.
CC -!- SUBUNIT: PH-DEPENDENT HOMODIMER OF ALPHA CHAINS AT PH VALUES LOWER
CC THAN 6.5 AND A HOMOTETRAMER OF ALPHA CHAINS ABOVE THIS VALUE.
CC -!- TISSUE SPECIFICITY: SEED.
CC -!- PTM: The beta and gamma chains are produced by partial proteolytic
CC processing of the lectin alpha chain by an asparaginyl
CC endopeptidase.

```

```

CC -!- MASS SPECTROMETRY: MW=25397; MW_ERR=3; METHOD=Electrospray;
CC RANGE=1-236.
CC -!- MASS SPECTROMETRY: MW=12847; MW_ERR=2; METHOD=Electrospray;
CC RANGE=1-118.
CC -!- MASS SPECTROMETRY: MW=12568; MW_ERR=2; METHOD=Electrospray;
CC RANGE=119-236.
CC -!- SIMILARITY: BELONGS TO THE LEGUMINOUS LECTIN FAMILY.
DR HSP; P02866; LDQ2.
DR InterPro; IPR000985; LECTIN_LEG.
DR InterPro; IPR001220; LECTIN_LEG.
DR Pfam; PF00138; LECTIN_LEG.
DR Pfam; PF00139; LECTIN_LEG.
DR PRODOM; PD000671; LECTIN_LEG.
DR PROSITE; PS000711; LECTIN_LEG.
DR PROSITE; PS00307; LECTIN_LEG.
DR PROSITE; PS00308; LECTIN_LEG.
KW LECTIN; Calcium; Manganese.
FT CHAIN 1 236 LECTIN ALPHA CHAIN.
FT CHAIN 1 118 LECTIN BETA CHAIN.
FT CHAIN 119 236 LECTIN GAMMA CHAIN.
SQ SEQUENCE 236 AA; 25398 MW; CA73BC77F4324251 CRC64;

```

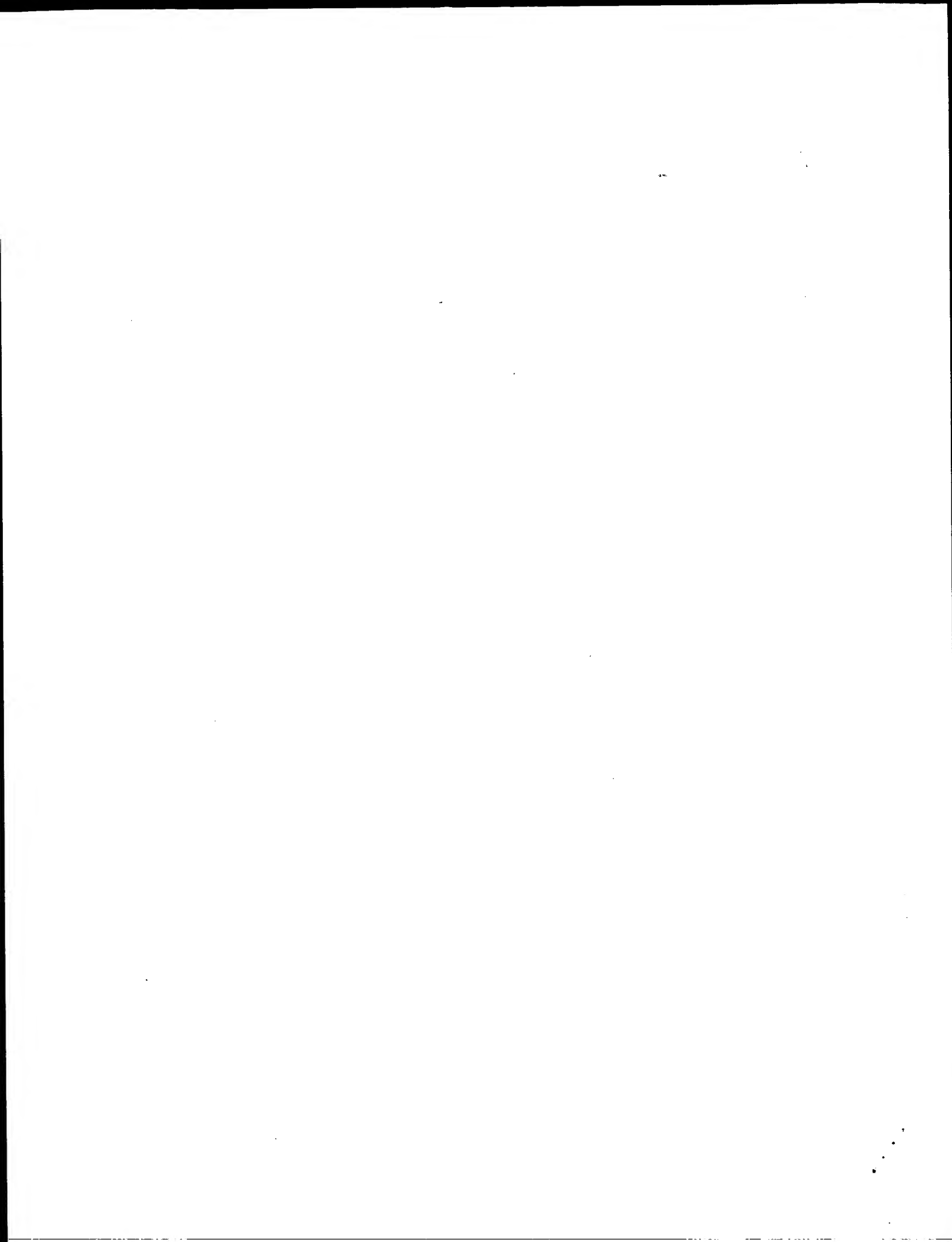
Query Match 53.08; Score 35; DB 1; Length 236;
 Best Local Similarity 63.68; Pred. No. 15;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 AQSLSFXFTKF 11
   | | | | |
DB 123 AOSLHFTFNOF 133

```

Search completed: February 26, 2003, 14:49:00
 Job time: 12 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:41 ; Search time 29 Seconds
(without alignments)
99.471 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SP_TREMBL_21.*
2: SP_Archea.*
3: SP_Bacteria.*
4: SP_Fungi.*
5: SP_Human.*
6: SP_Invertebrate.*
7: SP_Mhc.*
8: SP_Organelle.*
9: SP_Phage.*
10: SP_Plant.*
11: SP_Rodent.*
12: SP_Virus.*
13: SP_Vertebrate.*
14: SP_Unclassified.*
15: SP_Virus.*
16: SP_Bacteriap.*
17: SP_Archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	279	10 Q9M7M4	Q9M7M4 phaseolus v
2	55	83.3	272	10 Q9ZTA9	Q9ZTA9 dolichos la
3	42	63.6	408	16 Q9L1Y8	Q9L1Y8 pseudomonas
4	42	63.6	896	5 Q9VWV9	Q9VWV9 drosophila
5	42	63.6	928	5 Q9NIH2	Q9NIH2 drosophila
6	40	60.6	284	10 Q9FYU9	Q9FYU9 sophora fla
7	40	60.6	292	12 Q65239	Q65239 african swi
8	40	60.6	301	12 Q65196	Q65196 african swi
9	39	59.1	279	10 Q9ZWP4	Q9ZWP4 robinia pse
10	39	59.1	285	10 Q9ZWP5	Q9ZWP5 robinia pse
11	39	59.1	785	5 Q962M0	Q962M0 plasmodium
12	39	59.1	924	16 Q9I5U2	Q9I5U2 pseudomonas
13	38	57.6	254	10 Q43376	Q43376 arachis hyp
14	38	57.6	467	5 Q20265	Q20265 caenorhabdi
15	38	57.6	705	5 Q23024	Q23024 caenorhabdi
16	38	57.6	708	5 Q93380	Q93380 caenorhabdi

17	38	57.6	709	10 Q9FHY2	Q9FHY2 arabidopsis
18	38	57.6	807	5 Q18514	Q18514 caenorhabdi
19	38	57.6	2656	5 Q9GNU3	Q9GNU3 paracentrot
20	37	56.1	99	2 Q8RM68	Q8RM68 bacteroides
21	37	56.1	220	3 Q42911	Q42911 schizosacch
22	37	56.1	298	16 P73352	P73352 synchocyst
23	37	56.1	346	5 Q9N3W1	Q9N3W1 caenorhabdi
24	37	56.1	462	5 Q9TVQ1	Q9TVQ1 toxoplasma
25	36	54.5	129	12 Q11290	Q11290 molluscum c
26	36	54.5	284	5 Q17957	Q17957 caenorhabdi
27	36	54.5	318	16 Q8RDV1	Q8RDV1 fusobacteri
28	36	54.5	425	16 Q9A454	Q9A454 caulobacter
29	36	54.5	450	2 Q9EZ92	Q9EZ92 lactococcus
30	36	54.5	534	12 Q98195	Q98195 molluscum c
31	36	54.5	590	16 Q9RWT8	Q9RWT8 deinococcus
32	36	54.5	699	5 Q76608	Q76608 caenorhabdi
33	36	54.5	759	5 Q9SP54	Q9SP54 lymnaea sta
34	35	53.8	565	8 Q9T7K4	Q9T7K4 crassostrea
35	35	53.0	25	10 Q9S8C1	Q9S8C1 griffonia s
36	35	53.0	90	6 Q95LH3	Q95LH3 mustela put
37	35	53.0	110	4 Q75544	Q75544 homo sapien
38	35	53.0	124	3 Q07438	Q07438 saccharomyc
39	35	53.0	170	4 Q8WYJ4	Q8WYJ4 homo sapien
40	35	53.0	173	4 Q96LZ3	Q96LZ3 homo sapien
41	35	53.0	189	5 Q9VNF9	Q9VNF9 drosophila
42	35	53.0	191	6 Q95KJ8	Q95KJ8 macaca fasc
43	35	53.0	232	10 Q947I9	Q947I9 triticum tu
44	35	53.0	266	10 Q8RVX4	Q8RVX4 phaseolus v
45	35	53.0	270	10 Q40987	Q40987 pisum sativ

ALIGNMENTS

RESULT 1

Q9M7M4 ID Q9M7M4 PRELIMINARY; PRT; 279 AA.
AC Q9M7M4;
CT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mannose lectin FRIL (Fragment).
OS Phaseolus vulgaris (Kidney bean) (French bean).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Phaseolus.
OX NCBI_TaxID=3885;
RN [1]
RP SEQUENCE FROM N.A.
RA Moore J.G., Colucci G., Fuchs C.A., Hicklin D.J., Chrispeels M.J.,
RA Feldman M.;
RT "A new lectin in red kidney bean called PvFRIL stimulates
RT proliferation of NIH3T3 cells expressing the Flt3 receptor.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF121458; AAF28739.1;
DR HSSP; P02866; IONA.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; lectin_legA; 1.
DR Pfam; PF00139; lectin_legB; 1.
DR ProDom; PD000671; lectin_legA; 1.
DR ProDom; PD000711; lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
KW Lectin.
FT NON_TER
SQ SEQUENCE 279 AA; 31102 MW; F8919CF8B3EB4652 CRC64;

Query Match 97.0%; Score 64; DB 10; Length 279;
Best Local Similarity 92.9%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

Db 1 AQSLSFNTKFDLD 14

RESULT 2

Q9ZTA9 PRELIMINARY; PRT; 272 AA.
 AC Q9ZTA9; 1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-MAY-2002 (TREMBlrel. 20, Last annotation update)
 DE Mannoese lectin.
 GN FRIL.
 OS Dolichos lab lab (field bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
 OX NCBI_TaxID=35936;
 [1]
 RN RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=COTYLEDON;
 RC MEDLINE=99110944; PubMed=9892587;
 RA Colucci G., Moore J.G., Feldman M., Chrispeels M.J.;
 RT "cDNA cloning of FRIL, a lectin from Dolichos lablab, that preserves
 RT hemagglutinating properties in suspension culture.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:646-650(1999).
 DR EMBL: AF067417; AAD10734.1; -;
 DR HSSP: P02866; 1ONA.
 DR InterPro: IPR000985; Lectin_legA.
 DR InterPro: IPR001220; Lectin_legB.
 DR Pfam: PF00138; lectin_legA; 1.
 DR Pfam: PF00139; lectin_legB; 1.
 DR ProDom: PD000671; Lectin_legA; 1.
 DR ProDom: PD000711; Lectin_legB; 1.
 DR PROSITE: PS00308; LECTIN_LEGUME_ALPHA; 1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 DR Lectin.
 KW Lectin.
 SQ SEQUENCE 272 AA; 29900 MW; EA5C004307441495 CRC64;

Query Match 83.3%; Score 55; DB 10; Length 272;
 Best Local Similarity 91.7%; Pred. No. 0.0086;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFNTKFD 12
 Db 9 AQSLSFNTKFD 20

RESULT 3

Q911Y8 PRELIMINARY; PRT; 408 AA.
 AC Q911Y8;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Hypothetical protein PA2127.
 GN PA2127.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 [1]
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Broder R.L., Coulter L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Garber L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen.";

RL Nature 406:959-964(2000).
 DR EMBL: AE004640; AAG05515.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 408 AA; 47591 MW; 87FD5B3B704643B1 CRC64;

Query Match 63.6%; Score 42; DB 16; Length 408;
 Best Local Similarity 57.1%; Pred. No. 4.8;
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFNTKFDLD 14
 Db 385 ARSLSFQSQADID 398

RESULT 4

Q9VVV9 PRELIMINARY; PRT; 896 AA.
 AC Q9VVV9; Q9VVW0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Nkd protein.
 GN Nkd OR CG11614 OR CG18224.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Herman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003518; AAF49198.2; -;
 DR HSSP: P36610; 1G81.
 DR FlyBase: FBgn0002945; nkd.

SQ SEQUENCE 896 AA; 98892 MW; 3DC2AB6F07D246BB CRC64;
 Query Match 63.6%; Score 42; DB 5; Length 896;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AQSLSFXFTKFDLD 14
 : | | | | : | | |
 Db 158 SQPLQFSFTFYDLD 171
 : | | | | : | | |
 RESULT 5
 Q9NIH2 PRELIMINARY; PRT; 928 AA.
 ID Q9NIH2
 AC Q9NIH2
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Naked cuticle.
 GN NDK OR CG11614 OR CG18224.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Prorygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zeng W., Wharton K.A. Jr., Mack J.A., Wang K., Gadbaw M., Suyama K., Klein P.S., Scott M.P.;
 RT "Naked cuticle encodes an inducible antagonist of Wnt signalling.";
 RL Nature 0:0-0(2000).
 DR EMBL; AF213376; AAF34825.1; -.
 DR HSSP; P36610; IG81.
 DR FlyBase; FBgn0002945; nkd.
 SQ SEQUENCE 928 AA; 102591 MW; D543F378E2BC5122 CRC64;
 Query Match 63.6%; Score 42; DB 5; Length 928;
 Best Local Similarity 57.1%; Pred. No. 11;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AQSLSFXFTKFDLD 14
 : | | | | : | | |
 Db 190 SQPLQFSFTFYDLD 203
 : | | | | : | | |
 RESULT 6
 Q9FYU9 PRELIMINARY; PRT; 284 AA.
 ID Q9FYU9
 AC Q9FYU9
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Lectin.
 OS Sophora flavescens.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Sophoreae; Sophora.
 OX NCBI_TaxID=49840;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ROOT;
 RA Yan B., Ma Z.G., Wang L.X., Chai H.M., Huang X.Q.;
 RT "Cloning and sequencing of Sophora flavescens lectin gene, 284 AA.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF285121; AAG00508.1; -.
 DR HSSP; P02866; 1DQ2.
 DR InterPro; IPR000985; Lectin_legA.
 DR InterPro; IPR001220; Lectin_legB.
 DR Pfam; PF00138; lectin_legA; 1.
 DR Pfam; PF00139; lectin_legB; 1.
 DR ProDom; PD000671; Lectin_legA; 1.
 DR ProDom; PD000711; Lectin_legB; 1.
 DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; 1.

DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 SQ SEQUENCE 284 AA; 31318 MW; 2DC947EB3CBE0FB2 CRC64;
 Query Match 60.6%; Score 40; DB 10; Length 284;
 Best Local Similarity 66.7%; Pred. No. 8.3;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 AQSLSFXFTKFD 12
 : | | | | : | | |
 Db 31 ADSLSFTFSDFD 42
 : | | | | : | | |
 RESULT 7
 Q65239 PRELIMINARY; PRT; 292 AA.
 ID Q65239
 AC Q65239
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE ORF j15R.
 OS African swine fever virus (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
 OX NCBI_TaxID=10497;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MALAWI LIL20 /1;
 RX MEDLINE=94014996; PubMed=8409937;
 RA Vydelingum S., Baylis S.A., Bristow C., Smith G.L., Dixon L.K.;
 RT "Duplicated genes within the variable right end of the genome of a pathogenic isolate of African swine fever virus.";
 RL J. Gen. Virol. 74:2125-2130(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MALAWI LIL20 /1;
 RA Dixon L.K., Twigg S.R.F., Baylis S.A., Vydelingum S., Bristow C., Hammond J.M., Smith G.L.;
 RT "Nucleotide sequence of a 55 kbp region from the right end of the genome of a pathogenic African swine fever virus isolate (Malawi LIL20/1).";
 RL J. Gen. Virol. 7:1655-1684(1994).
 DR EMBL; X71982; CAA50835.1; -.
 SQ SEQUENCE 292 AA; 34549 MW; CA61A30E49D9219D CRC64;
 Query Match 60.6%; Score 40; DB 12; Length 292;
 Best Local Similarity 53.8%; Pred. No. 8.6;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 2 QSLSFXTKFDLD 14
 : | | | | : | | |
 Db 135 ETLTYFTDFDID 147
 : | | | | : | | |
 RESULT 8
 Q65196 PRELIMINARY; PRT; 301 AA.
 ID Q65196
 AC Q65196
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Proliferating cell nuclear antigen-like protein.
 GN E30LR.
 OS African swine fever virus (ASFV).
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
 OX NCBI_TaxID=10497;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA7IV;
 RX MEDLINE=21820291; PubMed=11831707;
 RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C., Rodriguez J.F., Vinuela E.;
 RT "Analysis of the complete nucleotide sequence of African swine fever virus.";
 RL Virology 208:249-278(1995).


```

RL J. Virol. 66:3860-3868(1992).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174976; PubMed=8438592;
RA Yanez R.J., Vinuela E.;
RT "African swine fever virus encodes a DNA ligase.";
RL Virology 193:531-536(1993).
RN [21]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174941; PubMed=8382399;
RA Pena L., Yanez R.J., Revilla Y., Vinuela E., Salas M.L.;
RT "African swine fever virus guanylyltransferase.";
RL Virology 193:319-328(1993).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=95159428; PubMed=7856088;
RA Simon-Mateo C., Freije J.M., Andres G., Lopez-Otin C., Vinuela E.;
RT "Mapping and sequence of the gene encoding protein p17, a major
RT African swine fever virus structural protein.";
RL Virology 206:1140-1144(1995).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato R., Freije J.M.P., Lopez-Otin C., Blasco R., Vinuela E.,
RA Salas M.L.;
RT "A gene homologous to topoisomerase II in african swine fever virus.";
RL Virology 188:938-947(1992).
RN [24]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94091056; PubMed=8266720;

Query Match 60.6%; Score 40; DB 12; Length 301;
Best Local Similarity 53.8%; Pred. No. 8.8;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QSLSPXFTKFDLD 14
Db 135 ETLEFFTFDFDID 147
      :| | | | | | | |

RESULT 9
Q92WP4 PRELIMINARY; PRT; 279 AA.
AC Q92WP4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Lectin-related polypeptide.
OS Robinia pseudoacacia (Black locust).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Robinieae; Robinia.
OX NCBI_TaxID=35938;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INNER BARK;
RA Yoshida K., Tazaki K.;
RT "Expression patterns of the genes that encode lectin or lectin-related
RT polypeptides in Robinia pseudoacacia.";
RL Aust. J. Plant Physiol. 26:495-502(1999).
DR EMBL; AB012635; BAA36416.1; -
DR HSSP; P02872; 2PEL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; Lectin_legA; 1.
DR Pfam; PF00139; Lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00308; LECTIN_LEGUME_ALPHA; UNKNOWN_1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
OX NCBI_TaxID=35938;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=INNER BARK;
RA Yoshida K., Tazaki K.;
RT "Expression patterns of the genes that encode lectin or lectin-related
RT polypeptides in Robinia pseudoacacia.";
RL Aust. J. Plant Physiol. 26:495-502(1999).
DR EMBL; AB012635; BAA36416.1; -
DR HSSP; P02872; 2PEL.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; Lectin_legA; 1.
DR Pfam; PF00139; Lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.

J. Virol. 66:3860-3868(1992).
RN [20]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174976; PubMed=8438592;
RA Yanez R.J., Vinuela E.;
RT "African swine fever virus encodes a DNA ligase.";
RL Virology 193:531-536(1993).
RN [21]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=93174941; PubMed=8382399;
RA Pena L., Yanez R.J., Revilla Y., Vinuela E., Salas M.L.;
RT "African swine fever virus guanylyltransferase.";
RL Virology 193:319-328(1993).
RN [22]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=95159428; PubMed=7856088;
RA Simon-Mateo C., Freije J.M., Andres G., Lopez-Otin C., Vinuela E.;
RT "Mapping and sequence of the gene encoding protein p17, a major
RT African swine fever virus structural protein.";
RL Virology 206:1140-1144(1995).
RN [23]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=92263807; PubMed=1316688;
RA Garcia-Beato R., Freije J.M.P., Lopez-Otin C., Blasco R., Vinuela E.,
RA Salas M.L.;
RT "A gene homologous to topoisomerase II in african swine fever virus.";
RL Virology 188:938-947(1992).
RN [24]
RP SEQUENCE FROM N.A.
RC STRAIN=BA71V;
RX MEDLINE=94091056; PubMed=8266720;

Query Match 59.1%; Score 39; DB 10; Length 285;
Best Local Similarity 70.0%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QSLSPXFTKTF 11
Db 35 ESVSPFTKTF 44
      :| | | | | | | |

RESULT 11
Q962M0 PRELIMINARY; PRT; 785 AA.
AC Q962M0;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE PVIH14070_F.
GN PVIH14070C.
OS Plasmodium vivax.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5655;
RN [1]
RP SEQUENCE FROM N.A.
RA Tchavtchitch M., Fischer K., Huestis R., Saul A.;
RT "The sequence of 200 kb portion of a Plasmodium vivax chromosome
RT reveals a high degree of conservation with P. falciparum chromosome
RT 3.";
RL Mol. Biochem. Parasitol. 0:0-0(2001).
DR EMBL; AY003872; AAF99460.1; -

```

```

DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR003960; AAA_sub.
DR Pfam; PF00004; AAA; 2.
DR PROSITE; PS00674; AAA; UNKNOWN_1.
SQ SEQUENCE 785 AA; 88003 MW; DB75ED37E13AFC63 CRC64;

Query Match
Best Local Similarity 59.1%; Score 39; DB 5; Length 785;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFK 11
   : : : : :
Db 591 AKSASFHTKF 601

RESULT 12
Q91502 PRELIMINARY; PRT; 924 AA.
AC Q91502;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Organic solvent tolerance protein OstA precursor.
GN OSTA OR PA0595.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizouchi S.D., Warren P.,
RA Stover C.K., Brinkman F.S.D., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy S., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong K.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE00495; AAG03984.1; -.
KW Complete proteome.
SQ SEQUENCE 924 AA; 104271 MW; 87F58C83E218335B CRC64;

Query Match
Best Local Similarity 59.1%; Score 39; DB 16; Length 924;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14
   : : : : :
Db 569 TLKYLTYKFDLD 580

RESULT 13
Q43376 PRELIMINARY; PRT; 254 AA.
AC Q43376;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mannose/glucose-binding lectin precursor (Fragment).
GN LEC.
OS Arachis hypogaea (Peanut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Aeschynomeneae;
OC Arachis.
OX NCBI_TaxID=3818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV- SELLIE; TISSUE=SEED;
RA Law I.J.;

*cloning and expression of cDNA for mannose-binding lectin from
peanut.*;
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RL EMBL; U22472; AAA74575.1; -.
DR HSSP; P02867; 2BQP.
DR InterPro; IPR000985; Lectin_legA.
DR InterPro; IPR001220; Lectin_legB.
DR Pfam; PF00138; Lectin_legA; 1.
DR Pfam; PF00139; Lectin_legB; 1.
DR ProDom; PD000671; Lectin_legA; 1.
DR ProDom; PD000711; Lectin_legB; 1.
DR PROSITE; PS00307; LECTIN_LEGWE_BETA; UNKNOWN_1.
KW Lectin; Signal.
FT SIGNAL 1
FT NON_TER 1
FT SIGNAL <1
FT SIGNAL 2
SQ SEQUENCE 254 AA; 28202 MW; 3AE1EA1F90B1CA03 CRC64;

Query Match
Best Local Similarity 57.6%; Score 38; DB 10; Length 254;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14
   : : : : :
Db 3 SLSFSYNKFEQD 14

RESULT 14
Q20265 PRELIMINARY; PRT; 467 AA.
ID Q20265;
AC Q20265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE F41C3.2 protein.
GN F41C3.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Willson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Chisoe S.;
RT "The sequence of C. elegans cosmid F41C3."
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL; U23521; AAC46809.1; -.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00083; sugar_tr; 1.
KW Transmembrane.
SQ SEQUENCE 467 AA; 51868 MW; D15F5181F0096C67 CRC64;

```

Query Match 57.6%; Score 38; DB 5; Length 467;
 Best Local Similarity 58.3%; Pred. No. 34;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDLD 14
 ||| |::| |
 Db 52 SLNFNFSKFEQD 63

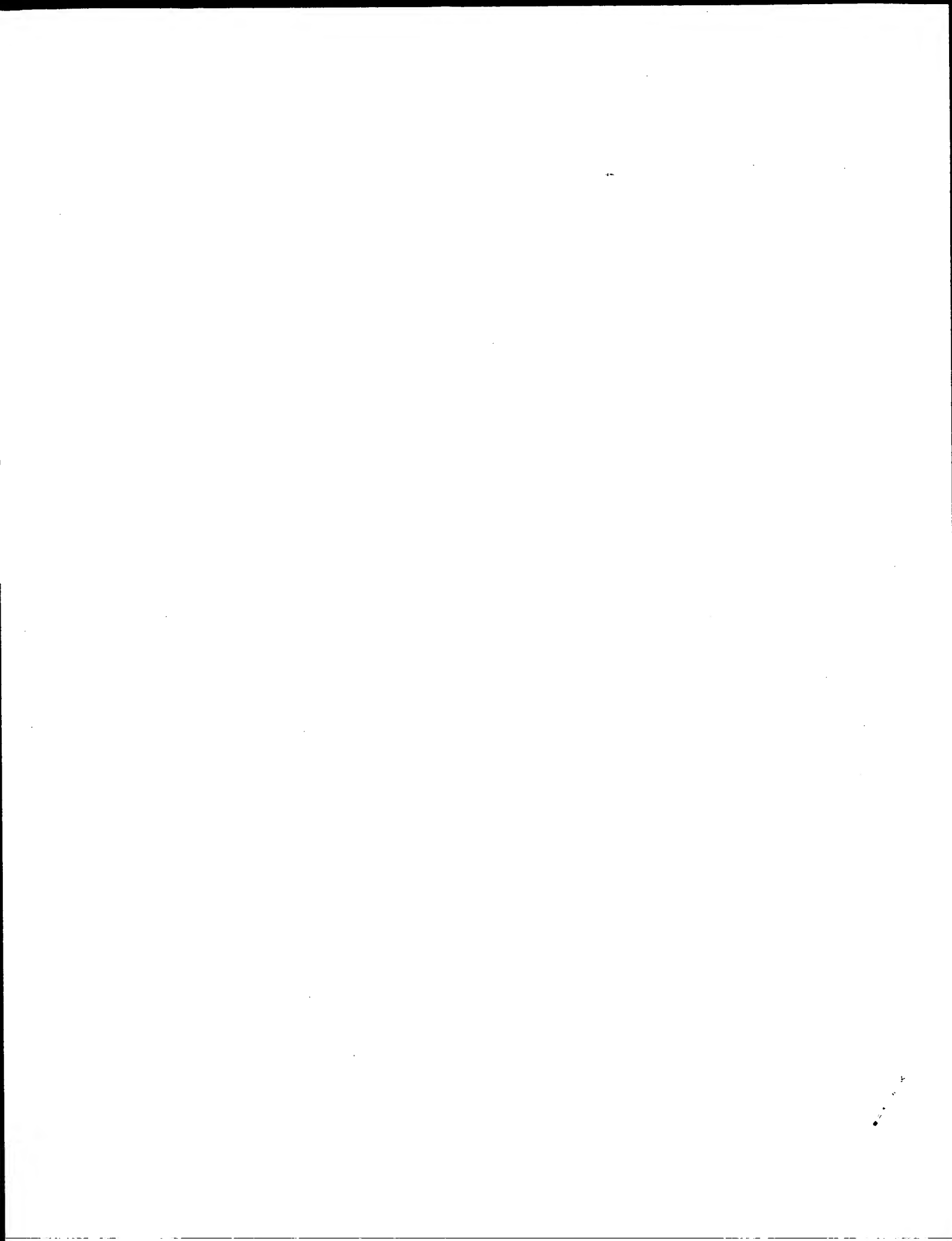
RESULT 15

Q23024 PRELIMINARY; PRT; 705 AA.
 AC Q23024; Q23025;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
 DE UNC-93 protein.
 GN UNC-93.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=92210636; PubMed=1313436;
 RA Levin J.Z., Horvitz R.H.;
 RT "The Caenorhabditis elegans unc-93 gene encodes a putative
 transmembrane protein that regulates muscle contraction.";
 RL J. Cell Biol. 117:143-155(1992).
 DR EMBL; X64415; CAA45760.1; -;
 DR EMBL; X64415; CAA45761.1; -;
 SQ SEQUENCE 705 AA; 80351 MW; 91D49A788A3EAC58 CRC64;

Query Match 57.6%; Score 38; DB 5; Length 705;
 Best Local Similarity 70.0%; Pred. No. 52;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 QSLSFXTKF 11
 ||| |::| |
 Db 623 QSLQFAFTKY 632

Search completed: February 26, 2003, 14:51:50
 Job time : 31 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:41 : Search time 83 Seconds
(without alignments)
22.476 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	14	AAW61498	Pylartin protein,
2	64	97.0	14	AAW67818	Flk2 ligand N-term
3	64	97.0	14	AAW62895	Antigenic peptide
4	64	97.0	303	AAW62898	Amino acid sequenc
5	56	84.8	15	AAW62899	Peptide derived fr
6	56	84.8	234	AAW62901	Amino acid sequenc
7	55	83.3	12	AAW61497	Pylartin protein,
8	55	83.3	264	AAW87973	A lectin derived p
9	55	83.3	264	AAW62890	Amino acid sequenc
10	55	83.3	286	AAW62894	Alpha-amylase inh

11	47	71.2	14	22	AAW62896	Peptide derived fr
12	42	63.6	896	22	ABB68966	Drosophila melanog
13	42	63.6	928	21	AAW08214	Amino acid sequenc
14	41	62.1	44	21	AAW62895	Arabidopsis thalia
15	41	62.1	44	21	AAW60788	Arabidopsis thalia
16	38	57.6	711	22	ABG08994	Human PRO polypept
17	37	56.1	73	22	AAU29221	Human PRO polypept
18	37	56.1	73	22	AAW87599	Human PRO3442. Ho
19	37	56.1	1108	21	AAW28530	Maize RNA-directed
20	35	53.0	87	22	AAW63706	Human gastric can
21	35	53.0	88	22	AAW84404	Human immune/haema
22	35	53.0	103	22	AAU31948	Novel human secret
23	35	53.0	139	22	ABW28896	Peptide #1547 enco
24	35	53.0	139	22	ABW34072	Peptide #1578 enco
25	35	53.0	139	22	ABW19510	Protein #1509 enco
26	35	53.0	139	22	AAW54854	Human brain expres
27	35	53.0	139	22	AAW15081	Peptide #1515 enco
28	35	53.0	139	22	AAW27529	Peptide #1566 enco
29	35	53.0	139	22	AAW02817	Peptide #1499 enco
30	35	53.0	139	22	ABW36886	Human peptide enco
31	35	53.0	170	21	AAW09977	Human CNBII protei
32	35	53.0	170	22	AAW14411	Calcineurin B subu
33	35	53.0	173	22	AAW64410	Amino acid sequenc
34	35	53.0	182	18	AAW20289	H. pylori transpor
35	35	53.0	187	22	AAW87327	Novel central nerv
36	35	53.0	189	22	ABW59336	Human testicular a
37	35	53.0	189	22	AAW87615	Novel central nerv
38	35	53.0	189	22	ABW58936	Drosophila melanog
39	35	53.0	189	22	AAW55239	Human reproductive
40	35	53.0	189	22	AAW43564	Human polypeptide
41	35	53.0	189	22	AAW43639	Human polypeptide
42	35	53.0	189	22	AAU19951	Novel human calciu
43	35	53.0	190	22	ABW71596	Drosophila melanog
44	35	53.0	193	22	AAW40369	Human polypeptide
45	35	53.0	193	22	AAW73635	Human S100 calcium

ALIGNMENTS

RESULT 1
AAW61498
ID AAW61498 standard; peptide; 14 AA.

XX AAW61498;

XX 19-OCT-1998 (first entry)

XX Pylartin protein, peptide chain beta (ii).

XX Pylartin protein; progenitor cell; haematopoietic system; cancer;
XX engraftation; haematologic disease; sickle cell anaemia; thalassemia.

OS Leguminosae.

XX WO9825457-A1.

XX 18-JUN-1998.

XX 09-DEC-1997; 97WO-US22486.

XX 28-MAR-1997; 97US-0825369.

XX 09-DEC-1996; 96US-0762537.

XX (INCL-) IMCLONE SYSTEMS INC.

XX Moore JG;

XX WPI; 1998-348161/30.

XX Pylartin proteins which preserve progenitor cells - useful for
PT haematopoietic therapies in cancer treatment or for treating
PT haematologic diseases

XX PS Claim 4; Page 28; 46pp; English.

CC CC The peptides AAW61497-W61502 can be used to form pylartin proteins which

CC CC preserve progenitor cells (pc). The proteins can be used for protecting

CC CC the integrity of the haematopoietic processes in vivo and as adjuncts in

CC CC therapeutic treatments related to cancer and other diseases which can

CC CC otherwise adversely impact upon the haematopoietic system. Since the

CC CC proteins bind specifically to primitive PCs, they can also be used for

CC CC the identification and localisation of PCs. The methods can be used for

CC CC e.g. expanding PC populations ex vivo to increase chances of

CC CC engraftation. Improving conditions for transporting and storing PCs and

CC CC for removing a fundamental barrier thereby enabling gene therapy to

CC CC treat and cure a broad range of life-threatening haematologic diseases

CC CC such as sickle cell anaemia and thalassemia.

XX SQ Sequence 14 AA;

Query Match 98.5%; Score 65; DB 19; Length 14;

Best Local Similarity 92.9%; Pred. No. 9.1e-06;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFSFTKFDLD 14

RESULT 2

AAAR67818

ID AAR67818 standard; peptide; 14 AA.

XX AC AAR67818;

XX DT 18-AUG-1995 (first entry)

XX DE Flk2 ligand N-terminal sequence.

XX KW human Flk2 receptor protein-tyrosine-kinase ligand; peripheral

XX KW peripheral blood leukocyte conditioned medium;

XX KW bone marrow disorder diagnosis; hematopoietic stem cell;

XX KW proliferation; differentiation.

XX OS Homo sapiens.

XX FH Key

FT misc_difference 7

FT /note= "any amino acid"

XX PN W09500554-A.

XX PD 05-JAN-1995.

XX PF 17-JUN-1994; 94WO-US06944.

XX PR 18-JUN-1993; 93US-0080244.

XX PR 21-JUN-1993; 93US-0081508.

XX PR 23-NOV-1993; 93US-0157490.

XX PA (UYPR-) UNIV PRINCETON.

XX PI Lemischka IR;

XX DR WPI; 1995-052014/07.

XX PT Ligand for receptor protein tyrosine kinase - useful for the

XX PT stimulation of primitive haematopoietic stem cells causing

XX PT proliferation and/or differentiation

XX PS Claim 1; Page 105; 131pp; English.

XX CC The sequence corresponds to the N-terminal region of a human Flk2

CC CC receptor protein-tyrosine-kinase ligand, isolated from

CC CC phytohemagglutinin-stimulated human peripheral blood leukocyte

CC tissue culture conditioned medium. The ligand may be used in

CC diagnosis of bone marrow disorders, and to stimulate the

CC proliferation and/or differentiation of primitive hematopoietic stem

CC cells. The ligand binds to a receptor protein-tyrosine-kinase

CC expressed in primitive but not mature mammalian hematopoietic cells.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 16; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFXFTKFDLD 14

RESULT 3

AAAG62895

ID AAG62895 standard; peptide; 14 AA.

XX AC AAG62895;

XX DT 17-SEP-2001 (first entry)

XX DE Antigenic peptide derived from a french bean FRIL polypeptide.

XX KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;

XX KW progenitor cell preservation factor; radiotherapy; chemotherapy;

XX KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

XX KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX OS Phaseolus vulgaris.

XX FH Key

FT misc_difference 7

FT /label= Asn, Cys, Ser

XX PN W0200149851-A1.

XX PD 12-JUL-2001.

XX PF 30-DEC-1999; 99WO-US31307.

XX PR 30-DEC-1999; 99WO-US31307.

XX PA (PHYL-) PHYLOGIX LLC.

XX PI Colucci MG, Chrispeels MJ, Moore JG;

XX DR WPI; 2001-441882/47.

XX PT Legume progenitor cell preservation factors for in vivo or ex vivo

XX PT preservation of hematopoietic progenitor cells and as therapeutics for

XX PT alleviating/reducing progenitor cell-depleting activity of cancer

XX PT therapeutics

XX PS Example 5; Page 72; 173pp; English.

XX CC The present sequence is derived from FRIL (Flk2/Flt3 tyrosine kinase

XX CC receptor-interacting lectin), and is used to raise antibodies. The

XX CC specification describes a composition of one or more members

XX CC of FRIL family of progenitor cell preservation factors. The composition

XX CC is useful for alleviating or reducing the hematopoietic progenitor

XX CC cell-depleting activity of a therapeutic treatment, including

XX CC radiotherapeutic and/or chemotherapeutic treatments. Administration of

XX CC FRIL compositions to a patient prior to treatment of the patient with

XX CC a therapeutic treatment having a hematopoietic progenitor cell-depleting

XX CC activity alleviates or reduces the hematopoietic progenitor

XX CC cell-depleting activity of the therapeutic treatment in the patient.

XX CC FRIL family members are useful for isolating population of progenitor

XX CC cells, hemangioblasts, and mesenchymal stem cells. The composition is

XX CC administered to reduce progenitor cell depleting effects of

CC chemotherapeutics, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX
 SQ Sequence 14 AA;
 Query Match 97.0%; Score 64; DB 22; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
 |||||
 DB 1 AQSLSFXFTKFDLD 14

RESULT 4
 AAG62898
 ID AAG62898 standard; Protein; 303 AA.

AC AAG62898;

DT 17-SEP-2001 (first entry)

XX Amino acid sequence of a french bean FRIL polypeptide.

DE FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX Phaseolus vulgaris.

OS WO200149851-A1.

PN 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

PI WPI; 2001-441882/47.

DR N-PSDB; AAH42306.

PT Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -

XX Example 5; Page 81; 173pp; English.

CC The present sequence represents a FRIL (FLK2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of
 CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing

CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 303 AA;

Query Match 97.0%; Score 64; DB 22; Length 303;
 Best Local Similarity 92.9%; Pred. No. 0.0004;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
 |||||
 DB 1 AQSLSFNFTKFDLD 14

RESULT 5

AAG62899

ID AAG62899 standard; peptide; 15 AA.

XX AAG62899;

DT 17-SEP-2001 (first entry)

DE Peptide derived from a yam FRIL polypeptide.

XX FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.

OS Sphenostylis stenocarpa.

XX WO200149851-A1.

PN 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.

XX 30-DEC-1999; 99WO-US31307.

XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;

PI WPI; 2001-441882/47.

XX Legume Progenitor cell preservation factors for in vivo or ex vivo
 XX preservation of hematopoietic progenitor cells and as therapeutics for
 XX alleviating/reducing progenitor cell-depleting activity of cancer
 XX therapeutics -

XX Example 22; Page 116; 173pp; English.

CC The present sequence is derived from a FRIL (FLK2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of a
 CC therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

DT 13-APR-1999 (first entry)
 XX A lectin derived progenitor cell preservation factor.
 DE
 XX
 XX Lactin derived progenitor cell preservation factor; progenitor cell;
 KW haematopoietic cell; cultured cell preservation; anticancer therapy;
 KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;
 KW FLK2/FLT3 receptor.
 XX
 XX Dolichos lab lab.
 OS
 XX
 XX WO9859038-A1.
 PN
 XX
 XX 30-DEC-1998.
 PD
 XX
 XX 23-JUN-1998; 98WO-US13046.
 PF
 XX
 XX 24-JUN-1997; 97US-0881189.
 PR
 XX
 XX (IMCL-) IMCLONE SYSTEMS INC.
 PA
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX
 XX Chrispeels MJ, Colucci MG, Moore JG;
 PI
 XX
 XX WPI; 1999-081274/07.
 DR
 XX N-PSDB; AAX03593.
 DR
 XX
 XX New nucleic acid encoding plant lectin that preserves progenitor
 PT cells - particularly haematopoietic progenitors, useful for bone
 PT marrow reconstitution after ablative therapy, and to increase DNA
 PT transfer in gene therapy
 PT
 XX
 PS Claim 1; Page 30-31; 72pp; English.
 XX
 CC The present sequence represents a lectin derived progenitor cell
 CC preservation factor. The protein is used to preserve unipotent,
 CC pluripotent or totipotent progenitor cells, especially haematopoietic
 CC cells, and also progenitors from nerve, muscle, skin, gut, bone,
 CC kidney, liver, pancreas or thymus. Specific applications are
 CC preservation of cultured cells intended for administration after
 CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation
 CC or chemotherapy) to reconstitute the haematopoietic system; enrichment
 CC of progenitor cells (e.g. during ex vivo purging of malignant cells);
 CC treatment of tissues containing haematopoietic progenitors for subsequent
 CC transplant to improve haematopoietic competence; improving transfer of
 CC exogenous DNA to progenitor cells (in gene therapy of various
 CC haematological disorders, e.g. sickle-cell anaemia); and protection
 CC against ablative therapy (to eliminate proliferating cells specifically),
 CC followed by re-establishment of differentiation and proliferation of
 CC preserved progenitors. The protein, when linked to magnetic beads, may
 CC also be used to isolate cells that express the FLK2/FLT3 receptor.
 XX
 SQ Sequence 264 AA;

Query Match 83.3%; Score 55; DB 20; Length 264;
 Best Local Similarity 91.7%; Pred. No. 0.017;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
 ||||| |||||
 Db 1 AQSLSFSFTKFD 12

RESULT 9
 AAG62890
 ID AAG62890 standard; Protein; 264 AA.
 XX
 AC AAG62890;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Amino acid sequence of a hyacinth bean FRIL polypeptide.
 XX

KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 XX severe combined immunodeficiency; aplastic anemia; tissue repair.
 OS Dolichos lab lab.
 XX
 XX WO200149851-A1.
 PN
 XX
 XX 12-JUL-2001.
 PD
 XX
 XX 30-DEC-1999; 99WO-US31307.
 PF
 XX
 XX 30-DEC-1999; 99WO-US31307.
 PR
 XX
 XX (PHYL-) PHYLOGIX LLC.
 PA
 XX Colucci MG, Chrispeels MJ, Moore JG;
 PI
 XX
 XX WPI; 2001-441882/47.
 DR
 XX N-PSDB; AAH42287.
 DR
 XX
 XX Legume progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics
 PT
 XX
 PS Example 1; Page 54-55; 173pp; English.
 XX
 CC The present sequence represents a FRIL (FLK2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of
 CC a therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of the chemotherapeutic, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.
 XX
 SQ Sequence 264 AA;

Query Match 83.3%; Score 55; DB 22; Length 264;
 Best Local Similarity 91.7%; Pred. No. 0.017;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
 ||||| |||||
 Db 1 AQSLSFSFTKFD 12

RESULT 10
 AAG62894
 ID AAG62894 standard; Protein; 286 AA.
 XX
 AC AAG62894;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Alpha-amylase inhibitor signal peptide and FRIL fusion.
 XX
 KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;
 KW progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 XX

KW severe combined immunodeficiency; aplastic anemia; tissue repair;
 KW alpha-amylase inhibitor gene.

XX Synthetic.
 OS Unidentified.
 OS Dolichos lab lab.
 XX WO200149851-A1.
 XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.
 XX 30-DEC-1999; 99WO-US31307.
 XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;
 PI WPI; 2001-441882/47.
 DR N-PSDB; AAH42295.
 XX Legume Progenitor cell preservation factors for in vivo or ex vivo
 PT preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -

XX Example 1; Page 59; 173pp; English.
 PS The present sequence represents fusion protein of alpha-amylase inhibitor
 CC signal peptide and FRIL (Flk2/Flt3 tyrosine kinase receptor-interacting
 CC lectin). The specification describes a composition of one or more members
 CC of FRIL family of progenitor cell preservation factors. The composition
 CC is useful for alleviating or reducing the hematopoietic progenitor
 CC cell-depleting activity of a therapeutic treatment, including
 CC radiotherapeutic and/or chemotherapeutic treatments. Administration of
 CC FRIL compositions to a patient prior to treatment of the patient with
 CC a therapeutic treatment having a hematopoietic progenitor cell-depleting
 CC activity alleviates or reduces the hematopoietic progenitor
 CC cell-depleting activity of the therapeutic treatment in the patient.
 CC FRIL family members are useful for isolating population of progenitor
 CC cells, hemangioblasts, and mesenchymal stem cells. The composition is
 CC administered to reduce progenitor cell depleting effects of
 CC chemotherapeutics, so that the patient can receive a higher dose of the
 CC chemotherapeutic and preferably recover from cancer. It is also
 CC administered to patients having, or predisposed to developing a
 CC condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 286 AA;
 SQ Query Match 83.3%; Score 55; DB 22; Length 286;
 Best Local Similarity 91.7%; Pred. No. 0.019;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12
 ||||| |||||
 Db 23 AQSLSFSFTKFD 34

RESULT 11
 AAG62896
 ID AAG62896 standard; peptide; 14 AA.
 XX AAG62896;
 AC AAG62896;
 XX 17-SEP-2001 (first entry)
 DT Peptide derived from a french bean FRIL polypeptide.
 DE FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
 XX progenitor cell preservation factor; radiotherapy; chemotherapy;
 KW

KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
 KW severe combined immunodeficiency; aplastic anemia; tissue repair.
 XX Phaseolus vulgaris.
 OS
 FH Key Location/Qualifiers
 FT Misc-difference 7 /note= "not known"
 XX WO200149851-A1.
 XX 12-JUL-2001.

XX 30-DEC-1999; 99WO-US31307.
 XX 30-DEC-1999; 99WO-US31307.
 XX (PHYL-) PHYLOGIX LLC.

XX Colucci MG, Chrispeels MJ, Moore JG;
 PI WPI; 2001-441882/47.
 DR Legume Progenitor cell preservation factors for in vivo or ex vivo
 XX preservation of hematopoietic progenitor cells and as therapeutics for
 PT alleviating/reducing progenitor cell-depleting activity of cancer
 PT therapeutics -

XX Example 5; Page 75; 173pp; English.
 PS The present sequence is derived from a FRIL (Flk2/Flt3 tyrosine kinase
 CC receptor-interacting lectin) polypeptide. The specification describes a
 CC composition of one or more members of FRIL family of progenitor cell
 CC preservation factors. The composition is useful for alleviating or
 CC reducing the hematopoietic progenitor cell-depleting activity of a
 CC therapeutic treatment, including radiotherapeutic and/or
 CC chemotherapeutic treatments. Administration of FRIL compositions to a
 CC patient prior to treatment of the patient with a therapeutic treatment
 CC having a hematopoietic progenitor cell-depleting activity alleviates or
 CC reduces the hematopoietic progenitor cell-depleting activity of the
 CC therapeutic treatment in the patient. FRIL family members are useful for
 CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
 CC stem cells. The composition is administered to reduce progenitor cell
 CC depleting effects of chemotherapeutics, so that the patient can receive
 CC a higher dose of the chemotherapeutic and preferably recover from cancer.
 CC It is also administered to patients having, or predisposed to developing
 CC a condition where the patients hematopoietic progenitor cells are
 CC depleted, such as severe combined immunodeficiency or aplastic anemia.
 CC The isolated mesenchymal cells are useful for tissue repair.

XX Sequence 14 AA;
 SQ Query Match 71.2%; Score 47; DB 22; Length 14;
 Best Local Similarity 85.7%; Pred. No. 0.024;
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14
 ||||| |||||
 Db 1 AQSLSFXTKDALD 14

RESULT 12
 ABB68966
 ID ABB68966 standard; Protein; 896 AA.
 XX ABB68966;
 AC ABB68966;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 33690.
 DE Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.
 KW

XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001WO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR WPI: 2001-656860/75.
 XX DR N-PSDB; ABL13069.
 XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX PS Disclosure; SEQ ID NO 33690; 2lpp + Sequence Listing; English.
 XX CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 896 AA;
 Query Match 63.6%; Score 42; DB 22; Length 896;
 Best Local Similarity 57.1%; Pred. No. 19;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AQSLSFXFTKFDLD 14
 :| | | | | :|||
 Db 158 SQPLQFSFTFYDLD 171
 RESULT 13
 AAB08214
 ID AAB08214 standard; Protein: 928 AA.
 XX AC AAB08214;
 XX DT 04-DEC-2000 (first entry)
 XX DE Amino acid sequence of a Drosophila naked cuticle polypeptide.
 XX KW Nkd gene; naked cuticle gene; segment-polarity gene; Wnt signalling;
 KW gene therapy; Nkd defect; cancer.
 XX OS Drosophila melanogaster.
 XX PN WO200049034-A1.
 XX PD 24-AUG-2000.
 XX PF 17-FEB-2000; 2000WO-US04198.
 XX PR 17-FEB-1999; 99US-0120646.
 XX PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX PR

PI Scott M, Zeng W, Wharton K;
 XX WPI: 2000-571967/53.
 DR N-PSDB; AAG63923.
 XX PT An isolated nucleic acid molecule useful for analyzing (genetic
 PT predisposition to) a disease state and for therapeutic purposes e.g.
 PT treatment of cancer comprises a sequence encoding a naked cuticle
 PT protein -
 XX PS Claim 12; Page 41-43; 58pp; English.
 XX CC The present sequence represents a Nkd (naked cuticle) polypeptide. In
 CC Drosophila, Nkd is a segment-polarity gene whose expression is
 CC induced by Wnt signalling. The Nkd polypeptide acts to antagonize Wnt
 CC signalling. Nkd may link ion fluxes to the regulation of Wnt signal
 CC potency, duration or distribution. The Nkd polynucleotides can be
 CC used for identifying homologous or related proteins, to modulate the
 CC expression or function of Nkd polypeptides, and in studying associated
 CC physiological pathways. Nkd polynucleotides can also be used in gene
 CC therapy to treat disorders associated with Nkd defects. They may
 CC also be used for therapeutic purposes e.g. treatment of cancer.
 XX SQ Sequence 928 AA;
 Query Match 63.6%; Score 42; DB 21; Length 928;
 Best Local Similarity 57.1%; Pred. No. 20;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AQSLSFXFTKFDLD 14
 :| | | | | :|||
 Db 190 SQPLQFSFTFYDLD 203
 RESULT 14
 AAG56205
 ID AAG56205 standard; Protein: 44 AA.
 XX AC AAG56205;
 XX DT 18-OCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 72209.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 29-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.


```

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.1%; Score 41; DB 21; Length 44;
Best Local Similarity 61.5%; Pred. NO. 1.1;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 OSLSEFTKEDLD 14
    :|||||  |||
Db 10 ESLSLFVGFD 22

RESULT 15
AAG60788
ID AAG60788 standard; Protein: 44 AA.
XX AC AAG60788;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 78778.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
XX PR 05-MAR-1999; 99US-0123180.
XX PR 09-MAR-1999; 99US-0123548.
XX PR 23-MAR-1999; 99US-0125788.
XX PR 25-MAR-1999; 99US-0126264.
XX PR 29-MAR-1999; 99US-0126785.
XX PR 01-APR-1999; 99US-0127482.
XX PR 06-APR-1999; 99US-0128234.
XX PR 08-APR-1999; 99US-0128714.
XX PR 16-APR-1999; 99US-0129845.
XX PR 19-APR-1999; 99US-0130077.
XX PR 21-APR-1999; 99US-0130449.
XX PR 23-APR-1999; 99US-0130510.
XX PR 23-APR-1999; 99US-0130891.
XX PR 28-APR-1999; 99US-0131449.
XX PR 30-APR-1999; 99US-0132048.
XX PR 30-APR-1999; 99US-0132407.
XX PR 04-MAY-1999; 99US-0132484.
XX PR 05-MAY-1999; 99US-0132485.
XX PR 06-MAY-1999; 99US-0132486.
XX PR 07-MAY-1999; 99US-0132487.
XX PR 11-MAY-1999; 99US-0132863.
XX PR 14-MAY-1999; 99US-0134218.
XX PR 14-MAY-1999; 99US-0134219.
XX PR 14-MAY-1999; 99US-0134221.
XX PR 14-MAY-1999; 99US-0134370.
XX PR 18-MAY-1999; 99US-0134768.
XX PR 19-MAY-1999; 99US-0134941.
XX PR 20-MAY-1999; 99US-0135124.
XX PR 21-MAY-1999; 99US-0135353.
XX PR 21-MAY-1999; 99US-0135629.
XX PR 25-MAY-1999; 99US-0136021.
XX PR 27-MAY-1999; 99US-0136392.
XX PR 28-MAY-1999; 99US-0136782.
XX PR 01-JUN-1999; 99US-0137222.
XX PR 03-JUN-1999; 99US-0137528.
XX PR 04-JUN-1999; 99US-0137502.
XX PR 07-JUN-1999; 99US-0137724.
XX PR 08-JUN-1999; 99US-0138094.
XX PR 10-JUN-1999; 99US-0138540.
XX PR 10-JUN-1999; 99US-0138847.
XX PR 14-JUN-1999; 99US-0139119.
XX PR 16-JUN-1999; 99US-0139452.
XX PR 16-JUN-1999; 99US-0139453.
XX PR 17-JUN-1999; 99US-0139454.
XX PR 18-JUN-1999; 99US-0139455.
XX PR 18-JUN-1999; 99US-0139456.
XX PR 18-JUN-1999; 99US-0139457.
XX PR 18-JUN-1999; 99US-0139458.
XX PR 18-JUN-1999; 99US-0139459.
XX PR 18-JUN-1999; 99US-0139460.
XX PR 18-JUN-1999; 99US-0139461.
XX PR 18-JUN-1999; 99US-0139462.
XX PR 18-JUN-1999; 99US-0139463.
XX PR 18-JUN-1999; 99US-0139750.
XX PR 18-JUN-1999; 99US-0139763.
XX PR 21-JUN-1999; 99US-0139817.
XX PR 22-JUN-1999; 99US-0139899.
XX PR 23-JUN-1999; 99US-0140353.
XX PR 23-JUN-1999; 99US-0140354.
XX PR 24-JUN-1999; 99US-0140695.
XX PR 28-JUN-1999; 99US-0140823.
XX PR 29-JUN-1999; 99US-0140991.
XX PR 30-JUN-1999; 99US-0141287.
XX PR 01-JUL-1999; 99US-0141842.
XX PR 01-JUL-1999; 99US-0142154.
XX PR 02-JUL-1999; 99US-0142055.
XX PR 06-JUL-1999; 99US-0142390.
XX PR 08-JUL-1999; 99US-0142803.
XX PR 09-JUL-1999; 99US-0142920.
XX PR 12-JUL-1999; 99US-0142977.
XX PR 13-JUL-1999; 99US-0143542.
XX PR 14-JUL-1999; 99US-0143624.
XX PR 15-JUL-1999; 99US-0144005.
XX PR 16-JUL-1999; 99US-0144085.
XX PR 16-JUL-1999; 99US-0144086.
XX PR 19-JUL-1999; 99US-0144325.
XX PR 19-JUL-1999; 99US-0144331.
XX PR 19-JUL-1999; 99US-0144332.
XX PR 19-JUL-1999; 99US-0144333.
XX PR 19-JUL-1999; 99US-0144334.
XX PR 19-JUL-1999; 99US-0144335.
XX PR 20-JUL-1999; 99US-0144352.
XX PR 20-JUL-1999; 99US-0144632.
XX PR 20-JUL-1999; 99US-0144884.
XX PR 21-JUL-1999; 99US-0144814.
XX PR 21-JUL-1999; 99US-0145086.
XX PR 21-JUL-1999; 99US-0145088.
XX PR 22-JUL-1999; 99US-0145085.
XX PR 22-JUL-1999; 99US-0145087.
XX PR 22-JUL-1999; 99US-0145089.
XX PR 22-JUL-1999; 99US-0145192.
XX PR 23-JUL-1999; 99US-0145145.
XX PR 23-JUL-1999; 99US-0145218.
XX PR 23-JUL-1999; 99US-0145224.

```

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 23-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 62.1%; Score 41; DB 21; Length 44;
Best Local Similarity 61.5%; Pred. No. 1.1;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSEXTKFDLD 14
 :|||||:||||
Db 10 ELSLFLFVGDFD 22

Search completed: February 26, 2003, 14:50:53
Job time : 84 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:00 ; Search time 12 Seconds
(without alignments)
44.009 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 174566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	14	10	US-09-934-251A-2
2	55	83.3	12	10	US-09-934-251A-1
3	42	63.6	60	10	US-09-730-989-22
4	42	63.6	668	10	US-09-730-989-20
5	37	56.1	73	9	US-10-063-547-148
6	37	56.1	73	9	US-10-174-590-396
7	37	56.1	73	9	US-10-176-758-396
8	37	56.1	73	9	US-10-063-616-148
9	37	56.1	73	9	US-10-175-737-396
10	37	56.1	73	9	US-10-063-502-148
11	37	56.1	73	9	US-10-173-706-396
12	37	56.1	73	9	US-10-175-738-396
13	37	56.1	73	9	US-10-175-752-396
14	37	56.1	73	9	US-10-176-482-396
15	37	56.1	73	9	US-10-176-757-396
16	37	56.1	73	9	US-10-176-913-396
17	37	56.1	73	9	US-10-180-552-396
18	37	56.1	73	9	US-10-180-557-396
19	37	56.1	73	9	US-10-173-700-396

20 37 56.1 73 9 US-10-174-572-396 Sequence 396, App
21 37 56.1 73 9 US-10-174-579-396 Sequence 396, App
22 37 56.1 73 9 US-10-174-582-396 Sequence 396, App
23 37 56.1 73 9 US-10-174-588-396 Sequence 396, App
24 37 56.1 73 9 US-10-175-739-396 Sequence 396, App
25 37 56.1 73 9 US-10-175-740-396 Sequence 396, App
26 37 56.1 73 9 US-10-175-743-396 Sequence 396, App
27 37 56.1 73 9 US-10-176-488-396 Sequence 396, App
28 37 56.1 73 9 US-10-176-492-396 Sequence 396, App
29 37 56.1 73 9 US-10-176-747-396 Sequence 396, App
30 37 56.1 73 9 US-10-176-750-396 Sequence 396, App
31 37 56.1 73 9 US-10-176-985-396 Sequence 396, App
32 37 56.1 73 9 US-10-176-987-396 Sequence 396, App
33 37 56.1 73 9 US-10-176-991-396 Sequence 396, App
34 37 56.1 73 9 US-10-176-992-396 Sequence 396, App
35 37 56.1 73 9 US-10-176-993-396 Sequence 396, App
36 37 56.1 73 9 US-10-184-658-396 Sequence 396, App
37 37 56.1 73 9 US-10-173-695-396 Sequence 396, App
38 37 56.1 73 9 US-10-173-697-396 Sequence 396, App
39 37 56.1 73 9 US-10-173-705-396 Sequence 396, App
40 37 56.1 73 9 US-10-174-576-396 Sequence 396, App
41 37 56.1 73 9 US-10-174-585-396 Sequence 396, App
42 37 56.1 73 9 US-10-174-586-396 Sequence 396, App
43 37 56.1 73 9 US-10-175-747-396 Sequence 396, App
44 37 56.1 73 9 US-10-176-481-396 Sequence 396, App
45 37 56.1 73 9 US-10-176-485-396 Sequence 396, App

ALIGNMENTS

RESULT 1
US-09-934-251A-2
; Sequence 2, Application US/09934251A
; Patent No. US20020132017A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934,251A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

Query Match 98.5%; Score 65; DB 10; Length 14;
Best Local Similarity 92.9%; Pred. No. 3.2e-06;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 AQSLSFXFTKFDLD 14
| | | | | | | | | | | | | | | |
Db 1 AQSLSFXFTKFDLD 14

RESULT 2
US-09-934-251A-1
; Sequence 1, Application US/09934251A
; Patent No. US20020132017A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934,251A

; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide chain of pylartin protein
US-09-934-251A-1

Query Match 83.3%; Score 55; DB 10; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDL 12
 |||||
DB 1 AQSLSFSFTKFD 12

RESULT 3
US-09-730-989-22
; Sequence 22, Application US/09730989
; Patent No. US20020061552A1
; GENERAL INFORMATION:
; APPLICANT: Yan, Dong
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: MAMMALIAN DISHEVELLED-ASSOCIATED PROTEINS
; FILE REFERENCE: PP-01657.002 / 200130.518
; CURRENT APPLICATION NUMBER: US/09/730,989
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-730-989-22

Query Match 63.6%; Score 42; DB 10; Length 60;
Best Local Similarity 57.1%; Pred. No. 0.26;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDL 14
 :| | | | :|||
DB 18 SLPLOFSFTFYDLD 31

RESULT 4
US-09-730-989-20
; Sequence 20, Application US/09730989
; Patent No. US20020061552A1
; GENERAL INFORMATION:
; APPLICANT: Yan, Dong
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: MAMMALIAN DISHEVELLED-ASSOCIATED PROTEINS
; FILE REFERENCE: PP-01657.002 / 200130.518
; CURRENT APPLICATION NUMBER: US/09/730,989
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-730-989-20

Query Match 63.6%; Score 42; DB 10; Length 668;
Best Local Similarity 57.1%; Pred. No. 3;

Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 AQSLSFXFTKFDL 14
 :| | | | :|||
DB 190 SQPLQFSFTFYDLD 203

RESULT 5
US-10-063-547-148
; Sequence 148, Application US/10063547
; Publication No. US20020182638A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,547
; CURRENT FILING DATE: 2002-05-02
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 148
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-547-148

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
 |||||
DB 47 SLSFYFLKFQL 57

RESULT 6
US-10-174-590-396
; Sequence 396, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-396

Query Match 56.1%; Score 37; DB 9; Length 73;

Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEXTKFDL 13
||||| |||
Db 47 SLSEYFLKFL 57

RESULT 7
US-10-176-758-396
; Sequence 396, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/176,758
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-758-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEXTKFDL 13
||||| |||
Db 47 SLSEYFLKFL 57

RESULT 8
US-10-063-616-148
; Sequence 148, Application US/10063616
; Publication No. US20030013855A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,616
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 148
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-616-148

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEXTKFDL 13
||||| |||
Db 47 SLSEYFLKFL 57

RESULT 9
US-10-175-737-396
; Sequence 396, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C104
; CURRENT APPLICATION NUMBER: US/10/175,737
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSEXTKFDL 13
||||| |||
Db 47 SLSEYFLKFL 57

RESULT 10
US-10-063-502-148
; Sequence 148, Application US/10063502
; Publication No. US20030023042A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,502
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 148
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-502-148

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKQQL 57

RESULT 11

US-10-173-706-396

; Sequence 396, Application US/10173706

; Publication No. US2003002293A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C7

; CURRENT APPLICATION NUMBER: US/10/173,706

; CURRENT FILING DATE: 2002-06-17

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 396

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-173-706-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKQQL 57

RESULT 12

US-10-175-738-396

; Sequence 396, Application US/10175738

; Publication No. US2003002294A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C45

; CURRENT APPLICATION NUMBER: US/10/175,738

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 396

; LENGTH: 73

; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-738-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKQQL 57

RESULT 13

US-10-175-752-396

; Sequence 396, Application US/10175752

; Publication No. US2003002295A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C60

; CURRENT APPLICATION NUMBER: US/10/175,752

; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 396

; LENGTH: 73

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-752-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKQQL 57

RESULT 14

US-10-176-482-396

; Sequence 396, Application US/10176482

; Publication No. US2003002296A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C70

; CURRENT APPLICATION NUMBER: US/10/176,482

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
| | | | | | | |
Db 47 SLSFYFLKFQL 57

RESULT 15

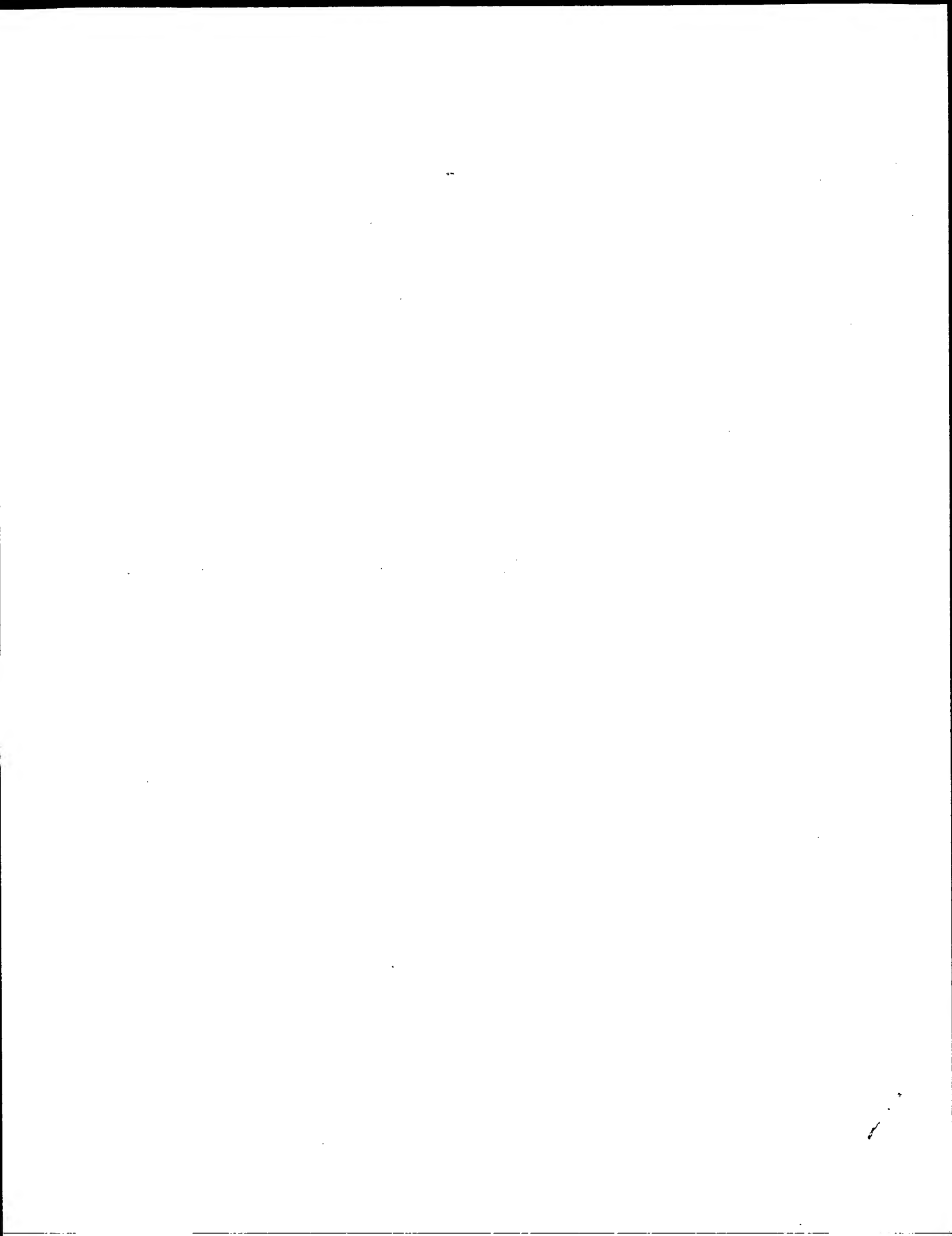
US-10-176-757-396

; Sequence 396, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C86
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-396

Query Match 56.1%; Score 37; DB 9; Length 73;
Best Local Similarity 72.7%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
| | | | | | | |
Db 47 SLSFYFLKFQL 57

Search completed: February 26, 2003, 14:55:42
Job time : 13 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:49:06 : Search time 139 Seconds
(without alignments)
64.937 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main.*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	14	1 PCT-US97-22486-2	Sequence 2, Appli
2	65	98.5	14	23 US-09-934-251A-2	Sequence 2, Appli
3	64	97.0	14	4 US-08-081-508-12	Sequence 12, Appl
4	64	97.0	14	5 US-08-157-490-11	Sequence 11, Appl
5	64	97.0	14	18 US-09-476-485A-31	Sequence 31, Appl
6	64	97.0	279	18 US-09-476-485A-56	Sequence 56, Appl

7	64	97.0	279	21	US-09-791-537-109055	Sequence 109055,
8	64	97.0	303	18	US-09-476-485A-6	Sequence 6, Appli
9	56	84.8	15	18	US-09-476-485A-9	Sequence 9, Appli
10	56	84.8	234	18	US-09-476-485A-8	Sequence 8, Appli
11	55	83.3	12	1	PCT-US97-22486-1	Sequence 1, Appli
12	55	83.3	12	23	US-09-934-251A-1	Sequence 1, Appli
13	55	83.3	105	1	PCT-US98-13046-13	Sequence 13, Appli
14	55	83.3	123	24	US-10-045-353-13	Sequence 13, Appli
15	55	83.3	123	18	US-09-476-485A-50	Sequence 50, Appli
16	55	83.3	237	18	US-09-476-485A-55	Sequence 55, Appli
17	55	83.3	270	1	PCT-US98-13046-2	Sequence 2, Appli
18	55	83.3	270	24	US-10-045-353-2	Sequence 2, Appli
19	55	83.3	272	21	US-09-791-537-1173	Sequence 1173, Ap
20	55	83.3	286	1	PCT-US98-13046-23	Sequence 23, Appli
21	55	83.3	286	18	US-09-476-485A-23	Sequence 23, Appli
22	55	83.3	286	24	US-10-045-353-23	Sequence 23, Appli
23	48	72.7	264	18	US-09-476-485A-2	Sequence 2, Appli
24	47	71.2	14	18	US-09-476-485A-32	Sequence 32, Appli
25	44.5	67.4	106	18	US-09-476-485A-49	Sequence 49, Appli
26	42	63.6	60	21	US-09-730-989-22	Sequence 22, Appli
27	42	63.6	195	16	US-09-270-767-61659	Sequence 61659, A
28	42	63.6	195	16	US-09-270-849B-193357	Sequence 193357, A
29	42	63.6	280	16	US-09-252-991A-16878	Sequence 16878, A
30	42	63.6	668	21	US-09-730-989-20	Sequence 20, Appli
31	42	63.6	687	16	US-09-270-767-46104	Sequence 46104, A
32	42	63.6	895	27	US-60-173-464-22125	Sequence 22125, A
33	42	63.6	896	20	US-09-614-150-33690	Sequence 33690, A
34	42	63.6	928	1	PCT-US00-04188-2	Sequence 2, Appli
35	42	63.6	928	19	US-09-506-066-2	Sequence 2, Appli
36	41	62.1	44	19	US-09-513-996A-72209	Sequence 72209, A
37	41	62.1	44	19	US-09-513-996A-78778	Sequence 78778, A
38	40	60.6	373	25	US-10-179-131-7061	Sequence 7061, Ap
39	39	59.1	279	21	US-09-791-537-115612	Sequence 115612, A
40	39	59.1	968	16	US-09-252-991A-17402	Sequence 17402, A
41	38.5	58.3	13	18	US-09-476-485A-34	Sequence 34, Appli
42	38	57.6	254	21	US-09-791-537-112588	Sequence 112588, A
43	38	57.6	458	21	US-09-733-089-18926	Sequence 18926, A
44	38	57.6	458	21	US-09-733-089-19024	Sequence 19024, A
45	38	57.6	458	22	US-09-816-660-18926	Sequence 18926, A

ALIGNMENTS

RESULT 1 22486-2
PCT-US97-22486-2
; Sequence 2, Application PC/TUS9722486
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR
; NUMBER OF INVENTIONS: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/22486
; FILING DATE: 9-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/762,537
; FILING DATE: 9-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: 08/825,369
; FILING DATE: 28-MAR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W. 37690
; REGISTRATION NUMBER: 381-21 CIP/PCT
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US97-22486-2

Query Match      98.5%; Score 65; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
DB 1 AQSLSFSFTKFDLD 14

RESULT 2
US-09-934-251A-2
; SEQUENCE 2, Application US/09934251A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934,251A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 14
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

Query Match      98.5%; Score 65; DB 23; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
DB 1 AQSLSFSFTKFDLD 14

RESULT 3
US-08-081-508-12
; SEQUENCE 12, Application US/08081508
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOPIPOTENT HEMATOPOIETIC STEM CELL
; TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
```

```
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/081,508
; FILING DATE: 19930621
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/975,049
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,941
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,272
; FILING DATE: 01-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 09-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 18-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N. 28,601
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-12P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-081-508-12

Query Match      97.0%; Score 64; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
DB 1 AQSLSFXFTKFDLD 14

RESULT 4
US-08-157-490-11
```

Sequence 11, Application US/08157490
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/157,490
FILING DATE: 23-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/975,049
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/005,941
FILING DATE: 15-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,272
FILING DATE: 01-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076022
FILING DATE: 09-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/080244
FILING DATE: 18-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/081508
FILING DATE: 21-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/096759
FILING DATE: 22-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125669
FILING DATE: 23-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-15P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-157-490-11

Query Match 97.0%; Score 64; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
Db 1 AQSLSFXFTKFDLD 14

RESULT 5

US-09-476-485A-31
Sequence 31, Application US/09476485A
GENERAL INFORMATION:
APPLICANT: Colucci, M. Gabriella
APPLICANT: Chrispeels, Maarten J.
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
TITLE OF INVENTION: and Products of Their Use
FILE REFERENCE: 108236.119
CURRENT APPLICATION NUMBER: US/09/476,485A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/881,189
PRIOR FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.0
SEQ ID NO 31
LENGTH: 14
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide corresponding to Pv-FRIL.
NAME/KEY: PEPTIDE
LOCATION: (7)..(7)
OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = Asn, Cys or Ser.
US-09-476-485A-31

Query Match 97.0%; Score 64; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
Db 1 AQSLSFXFTKFDLD 14

RESULT 6

US-09-476-485A-56
Sequence 56, Application US/09476485A
GENERAL INFORMATION:
APPLICANT: Colucci, M. Gabriella
APPLICANT: Chrispeels, Maarten J.
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
TITLE OF INVENTION: and Products of Their Use
FILE REFERENCE: 108236.119
CURRENT APPLICATION NUMBER: US/09/476,485A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/881,189
PRIOR FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn version 3.0
SEQ ID NO 56

; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PvFRIL.
US-09-476-485A-56

Query Match 97.0%; Score 64; DB 18; Length 279;
Best Local Similarity 92.9%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| |||||
Db 1 AQSLSFNFTKFDLD 14

RESULT 7
US-09-791-537-109055
; Sequence 109055, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 109055
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-09-791-537-109055

Query Match 97.0%; Score 64; DB 21; Length 279;
Best Local Similarity 92.9%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| |||||
Db 1 AQSLSFNFTKFDLD 14

RESULT 8
US-09-476-485A-6
; Sequence 6, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pv-FRIL.
US-09-476-485A-6

Query Match 97.0%; Score 64; DB 18; Length 303;
Best Local Similarity 92.9%; Pred. No. 0.0037;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| |||||
Db 1 AQSLSFNFTKFDLD 14

RESULT 9
US-09-476-485A-9
; Sequence 9, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Beta band polypeptide.
US-09-476-485A-9

Query Match 84.8%; Score 56; DB 18; Length 15;
Best Local Similarity 78.6%; Pred. No. 0.0051;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 10
US-09-476-485A-8
; Sequence 8, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: YamFril deduced amino acid sequence.
US-09-476-485A-8

Query Match 84.8%; Score 56; DB 18; Length 234;
Best Local Similarity 78.6%; Pred. No. 0.086; 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 11
PCT-US97-22486-1

```

; Sequence 1, Application PC/TUS9722486
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR
; TITLE OF INVENTION: PRESERVING PROGENITOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/22486
; FILING DATE: 9-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/762,537
; FILING DATE: 9-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/825,369
; FILING DATE: 28-MAR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US97-22486-1

```

```

Query Match      83.3%; Score 55; DB 1; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFD 12
    ||||| |||||
Db 1 AQSLSFSFTKFD 12

```

```

RESULT 12
US-09-934-251A-1
; Sequence 1, Application US/09934251A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934, 251A
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: peptide chain of pylartin protein
; US-09-934-251A-1
Query Match      83.3%; Score 55; DB 23; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFD 12
    ||||| |||||
Db 1 AQSLSFSFTKFD 12

```

```

RESULT 13
PCT-US98-13046-13
; Sequence 13, Application PC/TUS9813046
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: USA
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/13046
; FILING DATE: June 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: June 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US98-13046-13

```

```

Query Match      83.3%; Score 55; DB 1; Length 105;
Best Local Similarity 91.7%; Pred. No. 0.058;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 AQSLSFXFTKFD 12
    ||||| |||||
Db 1 AQSLSFSFTKFD 12

```

```

RESULT 14
US-10-045-353-13
; Sequence 13, Application US/10045353
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP

```

STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/045,353
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/881,189

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Feit, Irving N.

REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: 381-44 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 822-3550

TELEFAX: (516) 822-5582

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 13:

US-10-045-353-13

Query Match 83.3%; Score 55; DB 24; Length 105;
Best Local Similarity 91.7%; Pred. No. 0.058; 1; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12

Db 1 AQSLSFSFTKFD 12

RESULT 15

US-09-476-485A-50

; Sequence 50, Application US/09476485A

; GENERAL INFORMATION:

; APPLICANT: Colucci, M. Gabriella

; APPLICANT: Chrispeels, Maarten J.

; APPLICANT: Moore, Jeffrey G.

; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for

; FILE REFERENCE: 108236.119

; CURRENT APPLICATION NUMBER: US/09/476,485A

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: US 08/881,189

; PRIOR FILING DATE: 1997-06-24

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 50

; LENGTH: 123

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Beta-subunit of D1-FRIL.

US-09-476-485A-50

Query Match

Best Local Similarity

Matches 11; Conservative

0; Mismatches

1; Indels

0; Gaps

0;

83.3%; Score 55; DB 18; Length 123;

91.7%; Pred. No. 0.068;

0; Mismatches

1; Indels

0; Gaps

0;

QY 1 AQSLSFXFTKFD 12

Db 1 AQSLSFSFTKFD 12

Search completed: February 26, 2003, 14:54:16
Job time : 141 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:49:30 ; Search time 60 Seconds
(without alignments)
21.326 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 486122

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	97.0	279	6	US-10-190-258A-6
2	64	97.0	303	6	US-10-083-936B-6
3	56	84.8	15	6	US-10-083-936B-9
4	56	84.8	15	6	US-10-190-258A-9
5	56	84.8	234	6	US-10-083-936B-8
6	56	84.8	234	6	US-10-190-258A-8
7	55	83.3	264	6	US-10-083-936B-2
8	55	83.3	264	6	US-10-190-258A-2
9	37	56.1	73	6	US-10-125-923A-396
10	37	56.1	73	6	US-10-083-580-148
11	37	56.1	73	6	US-10-205-892-396
12	37	56.1	73	6	US-10-174-575-396
13	37	56.1	73	6	US-10-174-575A-396
14	37	56.1	73	6	US-10-187-755-396
15	37	56.1	73	6	US-10-187-749-396
16	37	56.1	73	6	US-10-139-672-396
17	37	56.1	73	6	US-10-063-557-148
18	37	56.1	73	6	US-10-194-486-396
19	37	56.1	1108	5	US-09-958-109A-6
20	35	53.0	139	6	US-10-203-138A-11864
21	35	53.0	168	5	US-09-724-676-71898
22	35	53.0	168	5	US-09-724-676-71900
23	35	53.0	168	5	US-09-724-676-71903
24	35	53.0	168	5	US-09-724-676-71904
25	35	53.0	168	5	US-09-724-676A-71898
26	35	53.0	168	5	US-09-724-676A-71900

27 35 53.0 168 5 US-09-724-676A-71903 Sequence 71903, A
28 35 53.0 168 5 US-09-724-676A-71904 Sequence 71904, A
29 35 53.0 191 5 US-09-724-676-71899 Sequence 71899, A
30 35 53.0 191 5 US-09-724-676-71902 Sequence 71902, A
31 35 53.0 191 5 US-09-724-676A-71899 Sequence 71899, A
32 35 53.0 191 5 US-09-724-676A-71902 Sequence 71902, A
33 35 53.0 193 1 PCT-US02-29560-268 Sequence 268, App
34 35 53.0 193 1 PCT-US02-29560-269 Sequence 269, App
35 35 53.0 193 5 US-09-724-676-48958 Sequence 48958, A
36 35 53.0 193 5 US-09-724-676-48959 Sequence 48959, A
37 35 53.0 193 5 US-09-724-676A-48958 Sequence 48958, A
38 35 53.0 193 5 US-09-724-676A-48959 Sequence 48959, A
39 35 53.0 193 6 US-10-245-882-268 Sequence 268, App
40 35 53.0 193 6 US-10-245-882-269 Sequence 269, App
41 35 53.0 329 6 US-10-050-902-217 Sequence 217, App
42 35 53.0 329 6 US-10-346-190-11 Sequence 11, Appl
43 35 53.0 329 6 US-10-050-898-217 Sequence 217, Appl
44 35 53.0 329 6 US-10-289-454-11 Sequence 11, Appl
45 35 53.0 625 6 US-10-092-411A-4504 Sequence 4504, Ap

ALIGNMENTS

RESULT 1
US-10-190-258A-6
; Sequence 6, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-190-258A-6

Query Match 97.0%; Score 64; DB 6; Length 279;
Best Local Similarity 92.9%; Pred. No. 2.1e-05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFKFDLD 14
| | | | | | | | | |
Db 1 AQSLSFNFTKFDLD 14

RESULT 2

US-10-083-936B-6
; Sequence 6, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/302,716
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Phaseolus vulgaris
US-10-083-936B-6

```

Query Match          97.0%; Score 64; DB 6; Length 303;
Best Local Similarity 92.9%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSLSFNFTKFDLD 14

RESULT 3
US-10-083-936B-9
; Sequence 9, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; PRIOR FILING DATE: 2002-10-22
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-083-936B-9

Query Match          84.8%; Score 56; DB 6; Length 15;
Best Local Similarity 78.6%; Pred. No. 3.1e-05;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 4
US-10-190-258A-9
; Sequence 9, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-9

Query Match          84.8%; Score 56; DB 6; Length 15;
Best Local Similarity 78.6%; Pred. No. 3.1e-05;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 5
US-10-083-936B-8
; Sequence 8, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; PRIOR FILING DATE: 2002-10-22
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-083-936B-8

Query Match          84.8%; Score 56; DB 6; Length 15;
Best Local Similarity 78.6%; Pred. No. 3.1e-05;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 6
US-10-190-258A-8
; Sequence 8, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-8

Query Match          84.8%; Score 56; DB 6; Length 234;
Best Local Similarity 78.6%; Pred. No. 0.00074;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 7
US-10-083-936B-2
; Sequence 2, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-9

Query Match          84.8%; Score 56; DB 6; Length 15;
Best Local Similarity 78.6%; Pred. No. 3.1e-05;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 8
US-10-083-936B-8
; Sequence 8, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-9

```

```

; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 60/302,716
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-083-936B-8

Query Match          84.8%; Score 56; DB 6; Length 234;
Best Local Similarity 78.6%; Pred. No. 0.00074;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 6
US-10-190-258A-8
; Sequence 8, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-8

Query Match          84.8%; Score 56; DB 6; Length 234;
Best Local Similarity 78.6%; Pred. No. 0.00074;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 7
US-10-083-936B-2
; Sequence 2, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-8

Query Match          84.8%; Score 56; DB 6; Length 234;
Best Local Similarity 78.6%; Pred. No. 0.00074;
Matches 11; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AQSLSFXFTKFDLD 14
   ||||| |||||
Db 1 AQSVSFTTKFDSD 14

RESULT 8
US-10-083-936B-8
; Sequence 8, Application US/10083936B
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Compositions and Methods for Protecting Tissues and
; FILE REFERENCE: 108236.130
; CURRENT APPLICATION NUMBER: US/10/083,936B
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: US 60/271,666
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-02-27
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Spheonostylis stenocarpa
US-10-190-258A-9

```


;
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-083-936B-2

Query Match 83.3%; Score 55; DB 6; Length 264;
Best Local Similarity 91.7%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
| | | | | | | | | |
Db 1 AQSLSFSFTKFD 12

RESULT 8
US-10-190-258A-2

; Sequence 2, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190.258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Dolichos lablab
US-10-190-258A-2

Query Match 83.3%; Score 55; DB 6; Length 264;
Best Local Similarity 91.7%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
| | | | | | | | | |
Db 1 AQSLSFSFTKFD 12

RESULT 9
US-10-125-923A-396

; Sequence 396, Application US/10125923A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C79
; CURRENT APPLICATION NUMBER: US/10/125.923A
; CURRENT FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 10/052586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24

Query Match 83.3%; Score 55; DB 6; Length 264;
Best Local Similarity 91.7%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
| | | | | | | | | |
Db 1 AQSLSFSFTKFD 12

; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-125-923A-396

Query Match 56.1%; Score 37; DB 6; Length 73;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 SLSFXFTKFDL 13
| | | | | | | | | |
Db 47 SLSFYFLKQFL 57

RESULT 10
US-10-063-580-148

; Sequence 148, Application US/10063580
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063.580
; CURRENT FILING DATE: 2002-05-03
; PRIOR APPLICATION NUMBER: 60/063435
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/064215
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 148
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-063-580-148

Query Match 56.1%; Score 37; DB 6; Length 73;
Best Local Similarity 72.7%; Pred. No. 1.5;

QY 3 SLSFXFTKFDL 13
| | | | | | | | | |
Db 47 SLSFYFLKQFL 57

; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

```

, TITLE OF INVENTION: ACIDS ENCODING THE SAME
, FILE REFERENCE: P3430RIC35
, CURRENT APPLICATION NUMBER: US/10/174,575
, CURRENT FILING DATE: 2002-06-18
, PRIOR APPLICATION NUMBER: 10/052586
, PRIOR FILING DATE: 2002-01-15
, PRIOR APPLICATION NUMBER: 60/059263
, PRIOR FILING DATE: 1997-09-18
, PRIOR APPLICATION NUMBER: 60/059266
, PRIOR FILING DATE: 1997-09-18
, PRIOR APPLICATION NUMBER: 60/062250
, PRIOR FILING DATE: 1997-10-17
, PRIOR APPLICATION NUMBER: 60/063120
, PRIOR FILING DATE: 1997-10-24
, PRIOR APPLICATION NUMBER: 60/063121
, PRIOR FILING DATE: 1997-10-24
, PRIOR APPLICATION NUMBER: 60/063486
, PRIOR FILING DATE: 1997-10-21
, PRIOR APPLICATION NUMBER: 60/063540
, PRIOR FILING DATE: 1997-10-28
, PRIOR APPLICATION NUMBER: 60/063541
, PRIOR FILING DATE: 1997-10-28
, PRIOR APPLICATION NUMBER: 60/063544
, PRIOR FILING DATE: 1997-10-28
, Remaining prior application data removed - See File Wrapper or PALM.

```

US-10-174-575A-396

Sequence 396. Application US/10174575A

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

APPLICANT: Watanabe, Colin K.

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-575A-396

Query Match 56.1%; Score 37; DB 6; Length 73;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKFQL 57

RESULT 14
US-10-187-755-396
; Sequence 396, Application US/10187755
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/187,755
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-755-396

Query Match 56.1%; Score 37; DB 6; Length 73;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKFQL 57

RESULT 15
US-10-187-749-396
; Sequence 396, Application US/10187749
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/187,749
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 396
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-187-749-396

Query Match 56.1%; Score 37; DB 6; Length 73;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 SLSFXFTKFDL 13
|||||
Db 47 SLSFYFLKFQL 57

Wed Feb 26 14:44:10 2003

us-09-476-485a-31.open.rapn

Page 6

Search completed: February 26, 2003, 14:55:22
Job time : 60 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:48:41 ; Search time 14 Seconds
(without alignments)
29.423 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	14	3	US-08-825-369A-2
2	55	83.3	12	3	US-08-825-369A-1
3	55	83.3	105	4	US-08-881-189B-13
4	55	83.3	270	4	US-08-881-189B-2
5	55	83.3	286	4	US-08-881-189B-23
6	35	53.0	191	3	US-08-655-352-5
7	35	53.0	191	3	US-08-655-352-6
8	35	53.0	191	4	US-09-258-016-5
9	35	53.0	191	4	US-09-258-016-6
10	35	53.0	191	4	US-09-257-825B-5
11	35	53.0	191	4	US-09-257-825B-6
12	35	53.0	625	4	US-09-134-001C-4504
13	34	51.5	187	1	US-07-956-700B-17
14	34	51.5	187	1	US-08-476-537-17
15	34	51.5	187	1	US-08-485-607-17
16	34	51.5	187	2	US-08-475-879-17
17	34	51.5	187	4	US-09-433-043B-17
18	34	51.5	191	3	US-08-655-352-7
19	34	51.5	191	4	US-09-258-016-7
20	34	51.5	191	4	US-09-257-825B-7
21	34	51.5	285	3	US-09-141-821-1
22	34	51.5	285	3	US-09-141-821-2
23	34	51.5	285	3	US-09-141-821-4
24	34	51.5	285	3	US-09-141-821-5
25	34	51.5	384	2	US-08-833-226-2
26	34	51.5	449	2	US-08-839-008-2
27	34	51.5	449	2	US-08-839-008-9

28 34 51.5 491 1 US-07-956-700B-109 Sequence 109, App
29 34 51.5 491 1 US-08-476-537-109 Sequence 109, App
30 34 51.5 491 1 US-08-485-607-109 Sequence 109, App
31 34 51.5 491 2 US-08-475-879-109 Sequence 109, App
32 34 51.5 491 4 US-09-433-043B-109 Sequence 109, App
33 34 51.5 491 4 US-09-433-043B-117 Sequence 117, App
34 34 51.5 559 1 US-08-424-788-3 Sequence 3, Appli
35 34 51.5 575 1 US-08-424-788-2 Sequence 2, Appli
36 34 51.5 575 1 US-08-110-683-4 Sequence 4, Appli
37 34 51.5 575 2 US-08-477-166-4 Sequence 4, Appli
38 34 51.5 575 2 US-08-472-097-4 Sequence 4, Appli
39 34 51.5 575 4 US-09-439-672-4 Sequence 4, Appli
40 34 51.5 575 5 PCT-US93-11638-4 Sequence 4, Appli
41 34 51.5 2172 1 US-08-611-107-31 Sequence 31, Appl
42 34 51.5 2257 1 US-08-611-107-10 Sequence 10, Appl
43 34 51.5 2257 2 US-08-422-560A-10 Sequence 10, Appl
44 34 51.5 2257 4 US-08-468-793-10 Sequence 10, Appl
45 33 50.0 169 4 US-08-720-625-4 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-825-369A-2
; Sequence 2, Application US/08825369A
; Patent No. 6084060
; GENERAL INFORMATION:
; APPLICANT: Moore
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/825.369A
; FILING DATE: March 28, 1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 381-21 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-825-369A-2

Query Match 98.5% Score 65; DB 3; Length 14;
Best Local Similarity 92.9%
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AQSLSFXFKFDLD 14

Db 1 AQSLSFXFKFDLD 14

RESULT 2

US-08-825-369A-1
; Sequence 1, Application US/08825369A
; Patent No. 6084060
; GENERAL INFORMATION:
; APPLICANT: Moore
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
; TITLE OF INVENTION: PROGENITOR CELLS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/825,369A
; FILING DATE: March 28, 1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 381-21 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-825-369A-1

Query Match 83.3%; Score 55; DB 3; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
DB 1 AQSLSFSFTKFD 12

RESULT 3
US-08-881-189B-13
; Sequence 13, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-13

Query Match 83.3%; Score 55; DB 4; Length 105;
Best Local Similarity 91.7%; Pred. No. 0.0015;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
DB 1 AQSLSFSFTKFD 12

RESULT 4
US-08-881-189B-2
; Sequence 2, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; TITLE OF INVENTION: PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 270 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-881-189B-2

Query Match 83.3%; Score 55; DB 4; Length 270;
Best Local Similarity 91.7%; Pred. No. 0.0042;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
 ||||| |||||
 Db 1 AQSLSFSFTKFD 12

RESULT 5
 US-08-881-189B-23
 ; Sequence 23, Application US/08881189B
 ; Patent No. 6310195
 ; GENERAL INFORMATION:
 ; APPLICANT: Colucci et al.
 ; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
 ; TITLE OF INVENTION: PROGNITOR CELL PRESERVATION FACTOR
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11753
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/881,189B
 ; FILING DATE: June 24, 1997
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Felt, Irving N.
 ; REGISTRATION NUMBER: 28,601
 ; REFERENCE/DOCKET NUMBER: 381-44 PCT
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 23:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 286 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; US-08-881-189B-23

Query Match 83.3%; Score 55; DB 4; Length 286;
 Best Local Similarity 91.7%; Pred. No. 0.0044;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
 ||||| |||||
 Db 23 AQSLSFSFTKFD 34

RESULT 6
 US-08-655-352-5
 ; Sequence 5, Application US/08655352
 ; Patent No. 6077991
 ; GENERAL INFORMATION:
 ; APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
 ; APPLICANT: Shameekumar Patil, Daisuke Takezawa
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
 ; ADDRESSEE: Whinston, LLP
 ; STREET: One World Trade Center
 ; STREET: 121 S.W. Salmon Street
 ; STREET: Suite 1600

CITY: Portland
 STATE: Oregon
 COUNTRY: United States of America
 ZIP: 97204
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Disk, 3-1/2 inch
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: MS DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/655,352
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/323,449
 ; FILING DATE: October 14, 1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dow, Alan. E.
 ; REGISTRATION NUMBER: 35,123
 ; REFERENCE/DOCKET NUMBER: 4630-45000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (503) 226-7391
 ; TELEFAX: (503) 228-9446
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 191 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; DESCRIPTION: Region of rat neural visinin-like protein
 ; DESCRIPTION: (Gen2:Ratnvp1) with homology to lily
 ; DESCRIPTION: CCaMK
 ; US-08-655-352-5

Query Match 53.0%; Score 35; DB 3; Length 191;
 Best Local Similarity 46.2%; Pred. No. 19;
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXFTKFD 14
 |:::|::|
 Db 99 QKLNWAFNMYDLD 111

RESULT 7
 US-08-655-352-6
 ; Sequence 6, Application US/08655352
 ; Patent No. 6077991
 ; GENERAL INFORMATION:
 ; APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
 ; APPLICANT: Shameekumar Patil, Daisuke Takezawa
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
 ; TITLE OF INVENTION: PRODUCTION OF MALE-STERILE PLANTS
 ; NUMBER OF SEQUENCES: 19
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Klarquist Sparkman Campbell Leigh &
 ; ADDRESSEE: Whinston, LLP
 ; STREET: One World Trade Center
 ; STREET: 121 S.W. Salmon Street
 ; STREET: Suite 1600
 ; CITY: Portland
 ; STATE: Oregon
 ; COUNTRY: United States of America
 ; ZIP: 97204
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Disk, 3-1/2 inch
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: MS DOS
 ; SOFTWARE: WordPerfect 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/655,352
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/323,449
FILING DATE: October 14, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan. E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 4630-45000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of chicken visinin-like protein
(Gen2:Ggvilp) with homology to lily
DESCRIPTION: CCaMK
US-08-655-352-6

Query Match 53.0%; Score 35; DB 3; Length 191;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLSPXFTKFDLD 14
| | | | :|||
Db 99 OKLNWAFNMYDLD 111

RESULT 8

US-09-258-016-5
Sequence 5, Application US/09258016
Patent No. 6362395

GENERAL INFORMATION:

APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
Shameekumar Patil, Daisuke Takezawa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
PRODUCTION OF MALE-STERILE PLANTS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
ADDRESS: Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,016
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stephens Jr., Donald L.
REGISTRATION NUMBER: 34,022
REFERENCE/DOCKET NUMBER: 4630-51994
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of rat neural visinin-like protein
(Gen2:Ratnvp1) with homology to lily

DESCRIPTION: CCaMK
US-09-258-016-5

Query Match 53.0%; Score 35; DB 4; Length 191;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLSPXFTKFDLD 14
| | | | :|||
Db 99 OKLNWAFNMYDLD 111

RESULT 9

US-09-258-016-6
Sequence 6, Application US/09258016
Patent No. 6362395

GENERAL INFORMATION:

APPLICANT: Bachettira W. Poovaiah, Zhihua Liu,
Shameekumar Patil, Daisuke Takezawa
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
PRODUCTION OF MALE-STERILE PLANTS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
ADDRESS: Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,016
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stephens Jr., Donald L.
REGISTRATION NUMBER: 34,022
REFERENCE/DOCKET NUMBER: 4630-51994
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of chicken visinin-like protein
(Gen2:Ggvilp) with homology to lily
DESCRIPTION: CCaMK
US-09-258-016-6

COMPUTER READABLE FORM:

MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/258,016
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Stephens Jr., Donald L.
REGISTRATION NUMBER: 34,022
REFERENCE/DOCKET NUMBER: 4630-51994
TELECOMMUNICATION INFORMATION:
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 191 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Region of chicken visinin-like protein
(Gen2:Ggvilp) with homology to lily
DESCRIPTION: CCaMK
US-09-258-016-6

Query Match 53.0%; Score 35; DB 4; Length 191;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLSPXFTKFDLD 14
| | | | :|||
Db 99 OKLNWAFNMYDLD 111

RESULT 10

US-09-257-825B-5
Sequence 5, Application US/09257825B
Patent No. 6403352

GENERAL INFORMATION:

APPLICANT: Poovaiah, Bachettira W.
APPLICANT: Patil, Shameekumar
APPLICANT: Takezawa, Daisuke
TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants
FILE REFERENCE: 4630-51993
CURRENT APPLICATION NUMBER: US/09/257,825B
CURRENT FILING DATE: 1999-02-25
PRIOR APPLICATION NUMBER: US 08/655,352
PRIOR FILING DATE: 1996-05-23
PRIOR APPLICATION NUMBER: US 60/014,743
PRIOR FILING DATE: 1996-03-28
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 191
TYPE: PRT
ORGANISM: Rat
US-09-257-825B-5

Query Match 53.0%; Score 35; DB 4; Length 191;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 0; Indels 4; Gaps 0;

QY 2 QSLSFXTKFDLD 14
| : : | : |||
Db 99 QKLNWAFNMYDLD 111

RESULT 11

US-09-257-825B-6
Sequence 6, Application US/09257825B
Patent No. 6403352
GENERAL INFORMATION:
APPLICANT: Poovaiah, Bachettira W.
APPLICANT: Patil, Shameekumar
APPLICANT: Takezawa, Daisuke
TITLE OF INVENTION: Compositions and Methods for Production of Male-Sterile Plants
FILE REFERENCE: 4630-51993
CURRENT APPLICATION NUMBER: US/09/257,825B
CURRENT FILING DATE: 1999-02-25
PRIOR APPLICATION NUMBER: US 08/655,352
PRIOR FILING DATE: 1996-05-23
PRIOR APPLICATION NUMBER: US 60/014,743
PRIOR FILING DATE: 1996-03-28
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 191
TYPE: PRT
ORGANISM: Chicken
US-09-257-825B-6

Query Match 53.0%; Score 35; DB 4; Length 191;
Best Local Similarity 46.2%; Pred. No. 19;
Matches 6; Conservative 3; Mismatches 0; Indels 4; Gaps 0;

QY 2 QSLSFXTKFDLD 14
| : : | : |||
Db 99 QKLNWAFNMYDLD 111

RESULT 12

US-09-134-001C-4504
Sequence 4504, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964

PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4504
LENGTH: 625
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4504

Query Match 53.0%; Score 35; DB 4; Length 625;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 LSFXTKFDL 13
| : : | : |||
Db 265 ISVHFNKFDL 274

RESULT 13

US-07-956-700B-17
Sequence 17, Application US/07956700B
Patent No. 5539092
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5539092th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
OPERATING SYSTEM: IBM PC Compatible
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/956,700B
FILING DATE: 19921002
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5539092thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-07-956-700B-17

Query Match 51.5%; Score 34; DB 1; Length 187;
Best Local Similarity 53.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
| : : | : |||
Db 149 KEISAVATKFDLD 161

RESULT 14

US-08-476-537-17
Sequence 17, Application US/08476537
Patent No. 5756290

GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5756290th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,537
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5756290thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-476-537-17

Query Match 51.5%; Score 34; DB 1; Length 187;
Best Local Similarity 53.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
Db 149 KEISAVATKFDLD 161

RESULT 15
US-08-485-607-17
Sequence 17, Application US/08485607
Patent No. 5792627
GENERAL INFORMATION:
APPLICANT: Robert Haselkorn and Piotr Gornicki
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA
TITLE OF INVENTION: Carboxylase
NUMBER OF SEQUENCES: 116
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: 321 No. 5792627th Clark Street
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60610
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,607
FILING DATE: 07-JUN-1995

CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/956,700
FILING DATE: 10/21/92
ATTORNEY/AGENT INFORMATION:
NAME: Thomas E. No. 5792627thrup
REGISTRATION NUMBER: 33,268
REFERENCE/DOCKET NUMBER: ARCD:058
TELECOMMUNICATION INFORMATION:
TELEPHONE: 1-312-744-0090
TELEFAX: 1-312-755-4489
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 187 amino acids
TYPE: Amino acid
STRANDEDNESS: Single
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
US-08-485-607-17

Query Match 51.5%; Score 34; DB 1; Length 187;
Best Local Similarity 53.8%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 QSLSFXTKFDLD 14
Db 149 KEISAVATKFDLD 161

Search completed: February 26, 2003, 14:51:13
Job time : 14 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:42:12 ; Search time 14 Seconds

(without alignments)
54.934 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 604

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	33.3	5	2 B31836	20K protein - Rick
2	12	33.3	7	2 S78024	ribosomal protein
3	12	33.3	8	2 S70727	ip9F protein - Shi
4	12	33.3	8	2 C61512	variant surface gl
5	12	33.3	8	2 A14683	aspartate transami
6	11	30.6	5	2 PT0295	Ig heavy chain CRD
7	11	30.6	6	2 S14159	parasporal crystal
8	11	30.6	6	2 B31263	dihydrofolate redu
9	11	30.6	7	2 S45311	microcin C7 - Esch
10	11	30.6	8	2 PL0184	capsid protein VP-
11	11	30.6	8	2 S68802	nitrate reductase
12	10	27.8	4	2 A37832	phenol 2-monooxyge
13	10	27.8	7	2 I40504	hypothetical prote
14	10	27.8	7	2 A34026	acetylcholinestera
15	10	27.8	8	2 S43971	tumor-associated a
16	10	27.8	8	2 S43972	tumor-associated a
17	10	27.8	8	2 A46306	spasmogenic toxin
18	10	27.8	8	2 A47618	beta-galactosidase
19	9	25.0	4	2 A48360	gamma subunit of p
20	9	25.0	5	2 A37114	hypoxanthine phosph
21	9	25.0	5	2 PT0625	T-cell receptor be
22	9	25.0	6	2 JH0784	neuropeptide TF-6
23	9	25.0	6	2 PT0560	T-cell receptor be
24	9	25.0	7	1 XEYDGD	galactose oxidase
25	9	25.0	7	2 S19630	ribosomal protein
26	9	25.0	7	2 B44787	callifMRamide 11
27	9	25.0	7	2 S68004	hucolin, 75K chain
28	9	25.0	7	2 PT0246	Ig heavy chain CRD
29	9	25.0	7	2 I48086	DNA topoisomerase

serine/threonine-s
polyphosphate-gluc
hypertrehalosemic
adipokinetic hormo
neuropeptide led-C
adipokinetic hormo
adipokinetic prote
hypothetical prote
protein QA300039 -
callifMRamide 8 -
cytochrome-c oxida
major postsynaptic
D-mannanate hydrol
branched-chain-ami
zinc-binding prote
Ig heavy chain CRD

ALIGNMENTS

RESULT 1

B31836

20K protein - Rickettsia rickettsii (fragment)

C:Species: Rickettsia rickettsii

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 28-May-1999

C:Accession: B31836

R:Anderson, B.E.; Baumstark, B.R.; Bellini, W.J.

J. Bacteriol. 170, 4493-4500, 1988

A:Title: Expression of the gene encoding the 17-kilodalton antigen from Rickettsia r

A:Reference number: A91885; MUID:89008059; PMID:3139629

A:Accession: B31836

A:Molecule type: DNA

A:Residues: 1-5 <AND>

A:Cross-references: GB:J03371; NID:gl52455; PIDN:AAI15030.1; PID:g4262874

Query Match

Best Local Similarity

Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3

Db 3 TNS 5

RESULT 2

S78024

ribosomal protein YmL21, mitochondrial - yeast (Saccharomyces cerevisiae) (fragment)

C:Species: Saccharomyces cerevisiae

C:Date: 09-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 14-Nov-1997

C:Accession: S78024

R:Kitakawa, M.; Graack, H.R.; Grohmann, L.; Goldschmidt-Reisin, S.; Herfurth, E.; Wit

Eur. J. Biochem. 245, 449-456, 1997

A:Title: Identification and characterization of the genes for mitochondrial ribosomal

A:Reference number: S78018; MUID:97296414; PMID:9151978

A:Accession: S78024

A:Molecule type: protein

A:Residues: 1-7 <KIT>

C:Genetics:

A:Genome: nuclear

C:Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match

Best Local Similarity

Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6

Db 4 NLK 7

RESULT 3

S70727

ipgF protein - Shigella flexneri (fragment)
 C:Species: Shigella flexneri
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 08-Oct-1999
 C:Accession: S70727
 R:Allaoui, A.; Sansonetti, P.J.; Menard, R.; Barzu, S.; Mounier, J.; Phalipon, A.; Parsc
 Mol. Microbiol. 17, 461-470, 1995
 A:Title: MxiG, a membrane protein required for secretion of Shigella spp. Ipa invasins:
 A:Reference number: S70727; MUID:96100445; PMID:8559065
 A:Accession: S70727
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-8 <ALL>
 A:Cross-references: EMBL:Z48957; NID:g929880; PIDN:CAA88821.1; PID:g9298801
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995
 C:Genetics:
 A:Gene: ipgF

Query Match 33.3%; Score 12; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3

||

Db 4 NN 5

RESULT 4

C61512
 Variant surface glycoprotein MITat 1.4 - Trypanosoma brucei (fragment)
 C:Species: Trypanosoma brucei
 C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-May-1999
 C:Accession: C61512
 R:Holder, A.A.; Cross, G.A.M.
 Mol. Biochem. Parasitol. 2, 135-150, 1981
 A:Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-termi
 A:Reference number: A61512; MUID:81172836; PMID:6163983
 A:Accession: C61512
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-8 <HOL>
 C:Keywords: glycoprotein

Query Match 33.3%; Score 12; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3

||

Db 3 NN 4

RESULT 5

AL4683
 aspartate transaminase (EC 2.6.1.1), mitochondrial - chicken (tentative sequence) (fragm
 N:Alternate names: aspartate aminotransferase, mitochondrial
 C:Species: Gallus gallus (chicken)
 C:Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 31-Mar-2000
 C:Accession: AL4683
 R:Wilson, K.J.; Hunziker, P.; Hughes, G.J.
 FEBS Lett. 108, 98-102, 1979
 A:Title: Microsequence analysis: IV. Automatic liquid-phase sequencing using DABITC.
 A:Reference number: A14683; MUID:80092116; PMID:520566
 A:Accession: AL4683
 A:Molecule type: protein
 A:Residues: 1-8 <WIL>
 C:Keywords: aminotransferase; mitochondrion

Query Match 33.3%; Score 12; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 3

||

Db 3 NN 4
 RESULT 6
 PT0295
 Ig heavy chain CRD3 region (clone 5-91) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
 R:Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
 C:Accession: PT0295
 J. Exp. Med. 173, 395-407, 1991
 A:Title: Preferential utilization of specific immunoglobulin heavy chain diversity an
 A:Reference number: PT0222; MUID:91108337; PMID:1899102
 A:Accession: PT0295
 A:Molecule type: DNA
 A:Residues: 1-5 <YAM>
 A:Experimental source: B lymphocyte
 C:Keywords: heterotetramer; immunoglobulin

Query Match 30.6%; Score 11; DB 2; Length 5;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3

||

Db 3 TRN 5

RESULT 7

SL4159
 Parasporal crystal protein CryIC - Bacillus thuringiensis (fragment)
 N:Alternate names: delta-endotoxin
 C:Species: Bacillus thuringiensis
 C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
 R:Convents, D.; Cherlet, M.; van Damme, J.; Lasters, I.; Lauwereys, M.
 Eur. J. Biochem. 195, 631-635, 1991
 A:Title: Two structural domains as a general fold of the toxic fragment of the Bacill
 A:Reference number: SL4087; MUID:91153300; PMID:1847865
 A:Accession: SL4159
 A:Molecule type: protein
 A:Residues: 1-6 <CON>

Query Match 30.6%; Score 11; DB 2; Length 6;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3

||

Db 3 TGN 5

RESULT 8

B31263
 dihydrofolate reductase (EC 1.5.1.3) / thymidylate synthase (EC 2.1.1.45) - Plasmodiu
 C:Species: Plasmodium falciparum
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 20-Mar-1996
 C:Accession: B31263
 R:Peterson, D.S.; Walliker, D.; Wellem, T.E.
 Proc. Natl. Acad. Sci. U.S.A. 85, 9114-9118, 1988
 A:Title: Evidence that a point mutation in dihydrofolate reductase-thymidylate syntha
 A:Reference number: A94217; MUID:89057886; PMID:2904149
 A:Accession: B31263
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-6 <PET>
 C:Keywords: methyltransferase; NADP; oxidoreductase

Query Match 30.6%; Score 11; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2.8e+05;
 Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2

Db 1 TN 2

RESULT 9

S45311
microcin C7 - Escherichia coli plasmid pmccc7
C:Species: Escherichia coli
C:Date: 10-Dec-1994 #sequence_revision 24-May-1996 #text_change 17-Mar-1999
C:Accession: S45311
R:Gonzalez-Pastor, J.E.; San Millan, J.L.; Moreno, F.
Nature 369, 281, 1994
A:Title: The smallest known gene.
A:Reference number: S45311; MUID:94239518; PMID:8183363
A:Accession: S45311
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-7 <GON>
C:Genetics:
A:Gene: mcca
A:Genome: plasmid pmccc7
C:Keywords: antibacterial
F:1/Modified site: N-formylmethionine #status predicted
F:7/Modified site: asparagine derivative (Asn) #status experimental

Query Match 30.6%; Score 11; DB 2; Length 7;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNW 3
I I
Db 3 TGN 5

RESULT 10

PL0184
capsid protein VP-1 - murine poliovirus (fragment)
C:Species: murine poliovirus, Theiler's encephalomyelitis virus
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
C:Accession: PL0184
R:Zurbriggen, A.; Hogle, J.M.; Fujinami, R.S.
J. Exp. Med. 170, 2037-2049, 1989
A:Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenicity
A:Reference number: PL0184; MUID:90063468; PMID:2479706
A:Accession: PL0184
A:Molecule type: genomic RNA
A:Residues: 1-8 <ZUR>
C:Keywords: capsid protein

Query Match 30.6%; Score 11; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
I I
Db 5 TN 6

RESULT 11

S68802
nitrate reductase (NADH) inhibitor - spinach (fragment)
C:Species: Spinacia oleracea (spinach)
C:Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C:Accession: S68802
R:Bachmann, M.; Huber, J.L.; Liao, P.C.; Gage, D.A.; Huber, S.C.
FEBS Lett. 387, 127-131, 1996
A:Title: The inhibitor protein of phosphorylated nitrate reductase from spinach (Spinacia)
A:Reference number: S68802; MUID:96244508; PMID:8674533
A:Accession: S68802
A:Molecule type: protein
A:Residues: 1-8 <BAC>
A:Experimental source: leaves; strain cv. Bloomsdale

Query Match 30.6%; Score 11; DB 2; Length 8;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVL 5
I I
Db 1 NLL 3

RESULT 12

A37832
phenol 2-monooxygenase (EC 1.14.13.7) chain P5 - Pseudomonas sp. (strain CF600) (frag
C:Species: Pseudomonas sp.
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Jun-1993
C:Accession: A37832
R:Powlowski, J.; Shingler, V.
J. Bacteriol. 172, 6834-6840, 1990
A:Title: In vitro analysis of polypeptide requirements of multicomponent phenol hydro
A:Reference number: A37832; MUID:91072231; PMID:2254259
A:Accession: A37832
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-4 <POW>
C:Keywords: oxidoreductase

Query Match 27.8%; Score 10; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NV 4
I I
Db 3 NV 4

RESULT 13

I40504
hypothetical protein 2 (7 aa) - Bacillus stearothermophilus
C:Species: Bacillus stearothermophilus
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 15-Oct-1999
C:Accession: I40504
R:Waye, M.M.; Winter, G.
Eur. J. Biochem. 158, 505-510, 1986
A:Title: A transcription terminator in the 5' non-coding region of the tyrosyl tRNA s
A:Reference number: I40503; MUID:86274732; PMID:3525162
A:Accession: I40504
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-7 <RES>
A:Cross-references: EMBL:X04193; MID:g40233; PIDN:CAA27782.1; PID:9580943

Query Match 27.8%; Score 10; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NV 4
I I
Db 3 NV 4

RESULT 14

A34026
acetylcholinesterase (EC 3.1.1.7) 5.6S form - Pacific electric ray (fragment)
C:Species: Torpedo californica (Pacific electric ray)
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 08-Nov-1996
C:Accession: A34026
R:Gibney, G.; MacPhee-Quigley, K.; Thompson, B.; Vedvick, T.; Low, M.G.; Taylor, S.S.
J. Biol. Chem. 263, 1140-1145, 1988
A:Title: Divergence in primary structure between the molecular forms of acetylcholine
A:Reference number: A34026; MUID:88087329; PMID:3355334
A:Accession: A34026
A:Molecule type: protein
A:Residues: 1-7 <GIB>
C:Keywords: alternative splicing; carboxylic ester hydrolase

Query Match 27.8%; Score 10; DB 2; Length 7;
Best Local Similarity 40.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VLQXT 8
: | |
Db 1 LLNAT 5

RESULT 15

S43971
tumor-associated antigen MMT1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Oct-1994 #sequence_revision 17-Nov-1995 #text_change 18-Aug-2000
C:Accession: S43971
R:Mandelboim, O.; Berke, G.; Fridkin, M.; Feldman, M.; Eisenstein, M.; Eisenbach, L.
Nature 369, 67-71, 1994
A:Title: CTL induction by a tumour-associated antigen octapeptide derived from a murine
A:Reference number: S43971; MUID:94217811; PMID:8164742
A:Accession: S43971
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <MAN>
C:Superfamily: unassigned animal peptides

Query Match 27.8%; Score 10; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 2; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NVLQ 6
: | |
Db 4 NTAQ 7

Search completed: February 26, 2003, 15:45:43
Job time : 15 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:39:16 ; Search time 11 seconds
(without alignments)
30.165 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 150

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	12	33.3	5	1 TPIS_CANFA	P54714 canis famil
2	12	33.3	8	1 FUSS_FUSSO	P81010 fusarium so
3	11	30.6	6	1 TMOF_SARBU	P41495 sarcophaga
4	11	30.6	6	1 VP19_HSVIK	P23210 herpes simp
5	11	30.6	8	1 FAR4_HOMAM	P41487 homarus ame
6	11	30.6	8	1 UPAL_HUMAN	P30087 homo sapien
7	10	27.8	7	1 FAR1_PROCL	P38499 procambarus
8	10	27.8	7	1 FAR2_PROCL	P38498 procambarus
9	10	27.8	8	1 FAR1_PANRE	P41872 panagrellus
10	10	27.8	8	1 FAR3_HOMAM	P41486 homarus ame
11	9	25.0	7	1 FAR6_CALVO	P41866 calliphora
12	9	25.0	7	1 IGAO_DAGDE	P06294 dactylum d
13	9	25.0	8	1 AKH_TABAT	P14595 tabanus atr
14	9	25.0	8	1 B44K_PORGI	P81886 porphyromon
15	9	25.0	8	1 COXC_RAT	P80430 rattus norv
16	9	25.0	8	1 FAR8_CALVO	P41863 calliphora
17	9	25.0	8	1 HTF2_PERAM	P04549 periplaneta
18	8	22.2	6	1 ACPH_RABIT	P25154 oryctolagus
19	8	22.2	7	1 FAR1_ASCSU	P31889 ascaris suu
20	8	22.2	7	1 FAR4_PANRE	P41875 panagrellus
21	8	22.2	7	1 MNPI_LEPDE	P42984 leptinotars
22	8	22.2	8	1 CAD1_ENTFA	P13268 enterococcu
23	8	22.2	8	1 CLP_THICU	P80488 thiobacillu
24	8	22.2	8	1 COW2_CONPU	P58785 conus purpu
25	7	19.4	5	1 BIOA_CITFR	P13071 citrobacter
26	7	19.4	5	1 TRM3_ECOLI	P13973 escherichia
27	7	19.4	6	1 ASP2_LACSN	P82655 lactobacill
28	7	19.4	6	1 UN06_CLOPA	P81351 clostridium
29	7	19.4	7	1 CHOX_ALCSP	P16101 alcaligenes
30	7	19.4	7	1 CIA_ENTFA	P11932 enterococcu
31	7	19.4	7	1 GFRE_MOUSE	P99025 mus musculu
32	7	19.4	7	1 LANC_CARUI	P36960 carnobacter
33	7	19.4	7	1 UN06_PTINP	P81675 pinus pinas

34	7	19.4	8	1 LCK1_LEUMA	P21140 leucophaea
35	7	19.4	8	1 LCK3_LEUMA	P21142 leucophaea
36	7	19.4	8	1 PLP_BRANA	P81707 brassica na
37	7	19.4	8	1 UPAA_HUMAN	P30096 homo sapien
38	6	16.7	4	1 FLRN_ANTEL	P58707 anthopleura
39	6	16.7	6	1 FARP_MONEX	P41966 moniezia ex
40	6	16.7	7	1 ALL5_CARMA	P81808 carcinus ma
41	6	16.7	7	1 CARP_MYTED	P10420 mytilus edu
42	6	16.7	7	1 UC24_MAIZE	P80630 zea mays (m
43	6	16.7	8	1 ACI_THUAL	P18691 thunnus alb
44	6	16.7	8	1 AKHG_GRYBI	P14086 gryllus bim
45	6	16.7	8	1 AKH_LIBAU	P25418 libellula a

ALIGNMENTS

RESULT 1					
TPIS_CANFA					
ID	TPIS_CANFA	STANDARD;	PRT;	5 AA.	
AC	P54714;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	15-JUL-1998 (Rel. 36, Last annotation update)				
DE	Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).				
GN	tpil.				
OS	Canis familiaris (Dog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
OX	NCBI_TaxID=9615;				
RN	[1]				
RP	SEQUENCE.				
KC	TISSUE=Heart;				
RA	MEDLINE=98163340; PubMed=9504812;				
RT	Dunn M.J., Corbett J.M., Wheeler C.H.;				
RT	"HSC-2DPAGE and the two-dimensional gel electrophoresis database of				
RL	dog heart proteins.";				
CC	Electrophoresis 18:2795-2802(1997).				
CC	!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate -> glycerone				
CC	phosphate.				
CC	!- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.				
CC	!- SUBUNIT: HOMODIMER (BY SIMILARITY).				
CC	!- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.				
DR	HSC-2DPAGE; P54714; DOG.				
DR	InterPro: IPR000652; Triophos_ismrse.				
DR	PROSITE; PS00171; TIM; PARTIAL.				
KW	Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;				
KW	Pentose shunt.				
FT	NON_TER 1				
FT	NON_TER 5				
SQ	SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;				
Query Match 33.3%; Score 12; DB 1; Length 5;					
Best Local Similarity 100.0%; Pred. No. 1.1e+05;					
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	2 NN 3				
Db	4 NN 5				
RESULT 2					
FUSS_FUSSO					
ID	FUSS_FUSSO	STANDARD;	PRT;	8 AA.	
AC	P81010;				
DT	15-JUL-1998 (Rel. 36, Created)				
DT	15-JUL-1998 (Rel. 36, Last sequence update)				
DT	15-JUL-1999 (Rel. 38, Last annotation update)				
DE	Allergen Fus s I3596* (Fragment).				
OS	Fusarium solani (subsp. pisi) (Nectria haematococca).				
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;				
OC	Hypocreales; Nectriaceae; Nectria.				
OX	NCBI_TaxID=70791;				

```

RN SEQUENCE.
RP STRAIN-IARI 3596; TISSUE=Myelium;
RA Verma J., Gangal S.V.;
RL Submitted (JUL-1997) to the SWISS-PROT data bank.
KW Allergen.
FT NON_TER
SQ SEQUENCE 8 AA; 898 MW; C372C441F5B69041 CRC64;

Query Match 33.3%; Score 12; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
DB 4 SHNV 7

RESULT 3
TMOF_SARBU STANDARD; PRT; 6 AA.
AC P41495;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Trypsin-modulating oostatic factor (TMOF).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Sarcophagidae; Sarcophaga.
OX NCBI_TaxID=7385;
RN [1]

SEQUENCE, AND SYNTHESIS.
RC TISSUE-Ovary;
RX MEDLINE=94211930; PubMed=8159807;
RA Byllemans D., Borovsky D., Hunt D.F., Shabanowitz J., Grauwels L.,
RA de Loof A.;
RT "Sequencing and characterization of trypsin modulating oostatic
RT factor (TMOF) from the ovaries of the grey fleshfly, Neobellieria
RT (Sarcophaga) bullata."
RL Regul. Pept. 50:61-72(1994).
CC -!- FUNCTION: HAS AN OOSTATIC ACTIVITY. INHIBITS TRYPSIN BIOSYNTHESIS
CC IN THE MIDGUT WHICH INDIRECTLY REDUCES THE VITELLOGENIN
CC CONCENTRATION IN THE HEMOLYMPH RESULTING IN INHIBITION OF OOCYTE
CC DEVELOPMENT.
CC -!- DEVELOPMENTAL STAGE: SYNTHESIZED AND RELEASED FROM FOLLICULAR
CC EPITHELIUM AFTER A BLOOD MEAL.
KW Hormone.
SQ SEQUENCE 6 AA; 695 MW; 61E72451B7642000 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 3 TN 4

RESULT 4
VP19_HSVIK STANDARD; PRT; 6 AA.
ID VP19_HSVIK
AC P23210;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Capsid assembly and DNA maturation protein (Virion protein UL38)
DE (Capsid protein VP19C) (Fragment).
GN UL38.
OS Herpes simplex virus (type 1 / strain KOS).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10306;

```

```

RN SEQUENCE FROM N.A.
RP MEDLINE=91101287; PubMed=1846198;
RA Flanagan W.M., Papavasiliou A.G., Rice M., Hecht L.B.,
RA Silverstein S., Wagner E.K.;
RT "Analysis of the herpes simplex virus type 1 promoter controlling the
RT expression of UL38, a true late gene involved in capsid assembly.";
RL J. Virol. 65:769-786(1991).
CC -!- FUNCTION: COMPONENT OF THE BASAL LAYER IN WHICH THE CAPSIDS ARE
CC EMBEDDED. BINDS DNA.
CC -!- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP19C.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57646; AAA45830.1; -
KW Capsid assembly; Coat protein; DNA-binding.
FT NON_TER
SQ SEQUENCE 6 AA; 703 MW; 67376451A336F000 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 3 TN 4

RESULT 5
FAR4_HOMAM STANDARD; PRT; 8 AA.
ID FAR4_HOMAM
AC P41487;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide 4 (FLI 4) (FL).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
SEQUENCE.
RC TISSUE-Pericardial organs;
RX MEDLINE=86116164; PubMed=3429714;
RA Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRFamide-like immunoreactive
RT substances from the lobster nervous system: isolation and sequence
RT analysis of two closely related peptides.";
RL J. Comp. Neurol. 266:16-26(1987)
CC -!- FUNCTION: CAN ACT AS A MODULATOR OF EXOSKELETAL AND CARDIAC
CC NEUROMUSCULAR JUNCTIONS.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES
SQ SEQUENCE 8 AA; 1067 MW; DDD40729C4540451 CRC64;

Query Match 30.6%; Score 11; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
DB 1 TN 2

```



```

RESULT 6
ID UPAL_HUMAN STANDARD; PRT; 8 AA.
AC P30087;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Unknown protein from 2D-page of plasma (Spot 2) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992)
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 4.9, ITS MW IS: 65 kDa.
DR SWISS-2DPAGE; P30087; HUMAN.
FT NON_TER 1
FT UNSURE 8
FT NON_TER 8
SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 30.6%; Score 11; DB 1; Length 8;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NN 4
   1
   1
Db 4 SN 6

RESULT 7
ID FAR1_PROCL STANDARD; PRT; 7 AA.
AC P38499;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cardioexcitatory FMRFamide homolog NFL.
OS Procamburus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacoidea; Astacoidea; Cambaridae; Procamburus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=93248032; PubMed=8387183;
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT "Isolation of two FMRFamide-related peptides from crayfish
RT pericardial organs.";
RL Peptides 14:137-143(1993)
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
   1
   1
Db 3 NFLR 6

RESULT 8
ID FAR2_PROCL STANDARD; PRT; 7 AA.
AC P38498;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Cardioexcitatory FMRFamide homolog DF2
OS Procamburus clarkii (Red swamp crayfish).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacoidea; Astacoidea; Cambaridae; Procamburus.
OX NCBI_TaxID=6728;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=93248032; PubMed=8387183;
RA Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT "Isolation of two FMRFamide-related peptides from crayfish
RT pericardial organs.";
RL Peptides 14:137-143(1993)
CC -!- FUNCTION: INCREASES THE RATE AND AMPLITUDE OF SPONTANEOUS
CC CONTRACTIONS OF SEMI-ISOLATED HEARTS. INCREASES THE AMPLITUDE OF
CC EXCITATORY POSTSYNAPTIC POTENTIALS IN ABDOMINAL EXTENSOR MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 967 MW; 69D40729C4540AC0 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 7;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
   1
   1
Db 3 NFLR 6

RESULT 9
ID FAR1_PANRE STANDARD; PRT; 8 AA.
AC P41872;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuropeptide PFI (SDPNFLRF-amide).
OS Panagrellus redivivus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX NCBI_TaxID=6233;
RN [1]
RP SEQUENCE.
RX MEDLINE=93027659; PubMed=1408999;
RA Geary T.G., Price D.A., Bowman J.W., Winterrowd C.A., Mackenzie C.D.,
RA Garrison R.D., Williams J.F., Friedman A.R.;
RT "Two FMRFamide-like peptides from the free-living nematode
RT Panagrellus redivivus.";
RL Peptides 13:209-214(1992).
CC -!- FUNCTION: MYOACTIVE.
CC -!- TISSUE SPECIFICITY: NERVE CORDS AND PAIRED GROUPS OF CELLS LOCATED
CC CAUDALLY TO THE BASE OF THE PHARYNX.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8
SQ SEQUENCE 8 AA; 995 MW; C6D40729C4576AB5 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 8;

```

Best Local Similarity 50.0%; Pred. No. 1.1e+05; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
| |
Db 4 NFLR 7

RESULT 10

FAR3_HOMAM
ID FAR3_HOMAM STANDARD; PRT; 8 AA.
AC P41486;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE FMRFamide-like neuro peptide 3 (FLI 3) (F2).
OS Homarus americanus (American lobster).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Astacidea; Nephropoidea; Nephropidae; Homarus.
OX NCBI_TaxID=6706;
RN [1]
RP SEQUENCE.
RC TISSUE=Pericardial organs;
RX MEDLINE=88116164; PubMed=3429714;
RA Trimmer B.A., Kobierski L.A., Kravitz E.A.;
RT "Purification and characterization of FMRFamide-like immunoreactive
RT substances from the lobster nervous system: isolation and sequence
RT analysis of two closely related peptides.";
RL J. Comp. Neurol. 266:16-26(1987).
CC -!- MISCELLANEOUS: PERICARDIAL ORGANS RELEASE THIS PEPTIDE WITH 100 MM
CC POTASSIUM IN THE PRESENCE OF CALCIUM.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1054 MW; C6D40729C4540AB5 CRC64;

Query Match 27.8%; Score 10; DB 1; Length 8;
Best Local Similarity 50.0%; Pred. No. 1.1e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
| |
Db 4 NFLR 7

RESULT 11

FARB_CALVO
ID FARB_CALVO STANDARD; PRT; 7 AA.
AC P41866;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Calliophora vomitoria 11.
OS Calliophora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnson A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated calliophorin) from the blowfly
RT Calliophora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
PR PIR; B44787; B44787.

KW Neuropeptide; Amidation.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 926 MW; 69D40699C44AB700 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 7;
Best Local Similarity 20.0%; Pred. No. 1.1e+05;
Matches 1; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 NVLQ 6
| |
Db 2 DNEMR 6

RESULT 12

IGAO_DACDE STANDARD; PRT; 7 AA.
ID IGAO_DACDE STANDARD; PRT; 7 AA.
AC P06294;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Galactose oxidase inhibitor.
OS Dactylium dendroides (Cladobotryum dendroides).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypomyces.
OX NCBI_TaxID=5132;
RN [1]
RP SEQUENCE.
RA Avigad G., Markus Z.;
RT "Identification of a peptide inhibitor of galactose oxidase from
RT Dactylium dendroides.";
RL Fed. Proc. 31:447-447(1972).
CC -!- FUNCTION: BINDS ONE COPPER ION PER MOLECULE BUT DOES NOT BIND THE
CC GALACTOSE OXIDASE APOENZYME. IT MAY INACTIVATE THE ENZYME BY
CC BINDING TO ITS PROSTHETIC COPPER GROUP.
DR PIR; A01341; XEYDGD.
KW Copper; Metalloenzyme inhibitor.
SQ SEQUENCE 7 AA; 706 MW; 75BB01A456D87DB0 CRC64;

Query Match 25.0%; Score 9; DB 1; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.1e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 OXT 8
| |
Db 3 QNT 5

RESULT 13

AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Adipokinetic hormone (AKH) (Diptera corpora cardiaca factor I)
DE (DCC 1).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Tabanomorpha; Tabanidae; Tabanus.
OX NCBI_TaxID=7207;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=90046758; PubMed=2813385;
RA Jarfe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT "Primary structure of two neuropeptide hormones with adipokinetic and
RT hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT flies (Diptera).";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF

CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
 CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
 CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
 DR PTR: A33995; A33995.
 DR InterPro: IPR002047; AKH.
 DR PROSITE: PS00256; AKH; 1.
 KW Neuropeptide; Amidation; Flight.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 949 MW; 86786771A9DIA736 CRC64;
 Query Match 25.0%; Score 9; DB 1; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 6 QXT 8
 Db 1 QLT 3
 RESULT 14
 B44K_PORGI STANDARD; PRT; 8 AA.
 AC P81886;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 44 kDa immunogenic protein (Fragment).
 OS Porphyromonas gingivalis (Bacteroides gingivalis).
 OC Bacteria; Bacteroidetes; Bacteroides; Bacteroidales;
 OC Porphyromonadaceae; Porphyromonas.
 OX NCBI_TaxID=837;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=VPB 3492;
 RA MEDLINE=20198497; PubMed=10731616;
 RX Norris J.M., Love D.N.;
 RT "Serum antibody responses of cats to soluble whole cell antigens of
 RT feline Porphyromonas gingivalis";
 RL Vet. Microbiol. 73:37-49(2000).
 CC -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
 KW Antigen.
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 989 MW; 9554540326CB476D CRC64;
 Query Match 25.0%; Score 9; DB 1; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1.1e+05;
 Matches 1; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 NV 4
 Db 7 NI 8
 RESULT 15
 COXG_RAT STANDARD; PRT; 8 AA.
 AC P80430;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytochrome c oxidase polypeptide VIB (EC 1.9.3.1) (AED) (Fragment).
 GN COX6B
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=95324529; PubMed=7601105;
 RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
 RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and

RT amino-terminal sequences suggest identity of the fetal heart and the
 RT adult liver isoform".
 RL Eur. J. Biochem. 230:235-241(1995).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT. THIS PROTEIN MAY BE ONE OF THE
 CC HEME-BINDING SUBUNITS OF THE OXIDASE.
 CC -!- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome
 CC c + 2 H(2)O.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIB FAMILY.
 KW Oxidoreductase; Mitochondrion.
 FT NON_TER 1 1
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 1039 MW; 8101E9CAA73AE456 CRC64;
 Query Match 25.0%; Score 9; DB 1; Length 8;
 Best Local Similarity 66.7%; Pred. No. 1.1e+05;
 Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 NVL 5
 Db 2 NXL 4
 Search completed: February 26, 2003, 15:43:43
 Job time : 12 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:41:42 : Search time 91 Seconds
(without alignments)
18,114 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 390

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	18	50.0	8	10 Q9SB24	Q9sb24 nicotiana t
2	17	47.2	8	2 O87471	O87471 haemophilus
3	15	41.7	4	11 Q08433	Q08433 rattus norv
4	15	41.7	8	4 Q15893	Q15893 homo sapien
5	14	38.9	8	2 Q8RJ10	Q8rj10 streptomyce
6	14	38.9	8	10 Q40659	Q40659 oryza sativ
7	13	36.1	8	3 Q05403	Q05403 saccharomyc
8	13	36.1	8	4 Q8TF70	Q8tf70 homo sapien
9	13	36.1	8	11 Q8R5M9	Q8r5m9 mus musculu
10	12	33.3	8	2 Q45615	Q45615 bacillus su
11	12	33.3	8	2 Q49534	Q49534 mycoplasma
12	11	30.6	7	2 Q47505	Q47505 escherichia
13	11	30.6	8	4 Q16468	Q16468 homo sapien
14	11	30.6	8	6 Q8WNS1	Q8wns1 bos taurus
15	11	30.6	8	8 Q9MSX1	Q9msx1 jurinea hum
16	11	30.6	8	11 Q62933	Q62933 rattus norv

17	11	30.6	8	12 Q9IU21	Q9iu21 influenza a
18	11	30.6	8	12 Q9IU19	Q9iu19 influenza a
19	11	30.6	8	13 P87488	P87488 oncorhynch
20	10	27.8	7	5 P83274	P83274 macrobrachi
21	10	27.8	7	12 Q9YQ10	Q9yq10 transmissib
22	10	27.8	8	2 Q9R9C2	Q9r9c2 borrelia bu
23	10	27.8	8	2 Q93SR0	Q93sr0 staphylococ
24	10	27.8	8	2 Q53790	Q53790 streptococc
25	10	27.8	8	4 Q9UMC7	Q9umc7 homo sapien
26	10	27.8	8	5 Q9UBI3	Q9ub13 albinaria h
27	10	27.8	8	5 P83275	P83275 macrobrachi
28	10	27.8	8	5 P83316	P83316 penaeus mon
29	10	27.8	8	8 Q9T2W0	Q9t2w0 saccharomyc
30	9	25.0	7	12 Q66113	Q66113 cherry leaf
31	9	25.0	7	12 Q67113	Q67113 influenzavi
32	9	25.0	8	2 Q56429	Q56429 thermus the
33	9	25.0	8	2 P72279	P72279 rhodococcus
34	9	25.0	8	2 Q51594	Q51594 escherichia
35	9	25.0	8	2 P77556	P77556 escherichia
36	9	25.0	8	4 Q15894	Q15894 homo sapien
37	9	25.0	8	4 Q15901	Q15901 homo sapien
38	9	25.0	8	4 Q9UDZ4	Q9udz4 homo sapien
39	9	25.0	8	4 Q96QF9	Q96qf9 homo sapien
40	9	25.0	8	8 Q9GB47	Q9gd47 hydraistele
41	9	25.0	8	8 Q9GCZ4	Q9gcz4 nanga punil
42	9	25.0	8	8 Q94VF6	Q94vf6 varanus job
43	9	25.0	8	8 Q8WGC9	Q8wgc9 upogebia af
44	9	25.0	8	8 Q8WGC7	Q8wgc7 petrollisthe
45	9	25.0	8	10 Q9SAY7	Q9say7 dioscorea t

ALIGNMENTS

RESULT 1					
Q9SB24					
ID Q9SB24	PRELIMINARY;	PRT;	8 AA.		
AC Q9SB24;					
DT 01-MAY-2000 (TREMBlrel. 13, Created)					
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)					
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)					
DE ParAt protein (Fragment).					
OS Nicotiana tabacum (Common tobacco).					
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.					
OX NCBI_TaxID=4097;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Niwa Y., Muranaka T., Baba A., Machida Y.;					
RT "Organ-specific and auxin-inducible expression of two tobacco para-					
RT related genes in transgenic plants.";					
RL DNA Res. 0:0-0(1994).					
DR EMBL; D42119; BAA07700.1; -.					
FT NON_TER	8				
SQ SEQUENCE	8 AA;	905 MW;	FE32D2C44455BB16 CRC64;		
Query Match	50.0%;	Score 18;	DB 10;	Length 8;	
Best Local Similarity	60.0%;	Pred. No. 6.7e+05;			
Matches	3;	Conservative	2;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1 TNNVL 5				
Db	3 SNNVV 7				
RESULT 2					
O87471					
ID O87471	PRELIMINARY;	PRT;	8 AA.		
AC O87471;					
DT 01-NOV-1998 (TREMBlrel. 08, Created)					
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)					
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)					

```

DE Hifa (Fragment).
GN HIFA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EAGAN;
RX MEDLINE=98389689; PubMed=9721313;
RA Mhlanga-Mutagadura T., Morlin G., Smith A.L., Eisenstark A.,
RA Golomb M.;
RT "Evolution of the major pilus gene cluster of haemophilus
RT influenzae.";
RL J. Bacteriol. 180:4693-4703(1998).
DR EMBL: AF071762; AAC35830.1; -.
FT NON_TER 1
SQ SEQUENCE 8 AA; 876 MW; DAB44451A7272325 CRC64;

Query Match 47.2%; Score 17; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
Db 5 TNN 7

RESULT 3
Q08433 PRELIMINARY; PRT; 4 AA.
ID Q08433
AC Q08433;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGT)
DE (Fragment).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Euteleostomi; Mammalia; Rodentia; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GUNN;
RX MEDLINE=91282758; PubMed=1840486;
RA Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT hyperbilirubinemic Gunn rat.";
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC ENDOGENOUS COMPOUNDS.
CC -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC -!- BETA-D-GLUCURONOSIDE.
CC EMBL: S38636; AAB19259.1; -.
KW transferase; Glycosyltransferase; Microsome; Multigene family.
FT NON_TER 1
FT NON_TER 4
SQ SEQUENCE 4 AA; 473 MW; 633732C420000000 CRC64;

Query Match 41.7%; Score 15; DB 11; Length 4;
Best Local Similarity 75.0%; Pred. No. 6.7e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NVLQ 6
Db 1 NVLK 4

RESULT 4
Q15893 PRELIMINARY; PRT; 8 AA.
ID Q15893

```

```

AC Q15893;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE (clone XP587A) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA Caskey C.T.H.;
RT "Isolation of chromosome-specific genes by reciprocal probing of
RT arrayed cDNAs and cosmid libraries.";
RL Hum. Mol. Genet. 0:0-0(1995).
DR EMBL: L32073; AAA73883.1; -.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 874 MW; DAA1B6D7376456C5 CRC64;

Query Match 41.7%; Score 15; DB 4; Length 8;
Best Local Similarity 37.5%; Pred. No. 6.7e+05;
Matches 3; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8
Db 1 SQNPLOYS 8

RESULT 5
Q8RJ10 PRELIMINARY; PRT; 8 AA.
ID Q8RJ10
AC Q8RJ10;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA replication initiation protein (Fragment).
GN DNA-LIKE.
OS Streptomyces caespitosus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=53502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27422;
RA Ma W., Mao X., Jie L., Qin Z.J., Jiang W.H., Chiao R.S., Zhao G.P.;
RT "Cloning, sequence analysis and function analysis of the replication
RT origin oric of Streptovorticillum caespitosus ATCC27422.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ458440; CAD30324.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 8 AA; 915 MW; 04744330440451A7 CRC64;

Query Match 38.9%; Score 14; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 2 TNRI 5

RESULT 6
Q40659 PRELIMINARY; PRT; 8 AA.
ID Q40659
AC Q40659;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Alpha-amylase (Fragment).

```

```

OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91078641; PubMed=2258052;
RA Kumagai M.H., Shah M., Terashima M., Vrkljan Z., Whitaker J.R.,
RA Rodriguez R.L.;
RT "Expression and secretion of rice alpha-amylase by saccharomycetes
RT cerevisiae.";
RL Gene 94:209-216(1990).
DR EMBL; M62916; AAA33892.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 948 MW; EBC694444732D6D6 CRC64;

Query Match 38.9%; Score 14; DB 10; Length 8;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVL 5
Db 5 NMV 8

RESULT 7
Q05403 ID Q05403 PRELIMINARY; PRT; 8 AA.
AC Q05403;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA for ORF's from chromosome XV (Fragment).
GN COQ3 AND YOL096C.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=FV1679;
RA Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT "A 29,425 kb segment on the left arm of yeast chromosome XV contains
RT more than twice as many unknown as known open reading frames.";
RL Yeast 11:975-986(1995).
DR EMBL; X83121; CAA58183.1; -.
DR SGD; S0005456; COQ3.
FT NON_TER 8
SQ SEQUENCE 8 AA; 879 MW; 7B5322D2C441E058 CRC64;

Query Match 36.1%; Score 13; DB 3; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6
Db 3 HNVVK 7

RESULT 8
Q8TF70 ID Q8TF70 PRELIMINARY; PRT; 8 AA.
AC Q8TF70;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microphthalma-associated transcription factor (Fragment).
GN MITF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takeda K., Yasumoto K., Shibahara S.;
RT "An Mitf isoform exclusively expressed in the affected cells of Mitf
RT mutants.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061771; BAB85121.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 989 MW; ED0727204415A1A6 CRC64;

Query Match 36.1%; Score 13; DB 4; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 2 TSRIL 6

RESULT 9
Q8R5M9 ID Q8R5M9 PRELIMINARY; PRT; 8 AA.
AC Q8R5M9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Microphthalma-associated transcription factor (Fragment).
GN MITF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=EYE;
RA Takeda K., Yasumoto K., Shibahara S.;
RT "An Mitf isoform exclusively expressed in the affected cells of Mitf
RT mutants.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB061769; BAB85120.1; -.
FT NON_TER 8
SQ SEQUENCE 8 AA; 989 MW; ED0727204415A1A6 CRC64;

Query Match 36.1%; Score 13; DB 11; Length 8;
Best Local Similarity 40.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5
Db 2 TSRIL 6

RESULT 10
Q45615 ID Q45615 PRELIMINARY; PRT; 8 AA.
AC Q45615;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE GutB protein (Fragment).
GN GUTB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RA Ye R., Wong S.L.;
RT "Transcriptional regulation of the Bacillus subtilis glucitol
RT dehydrogenase gene.";

```

Best Local Similarity 66.7%; Pred No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3
|
|
Db 3 TGN 5

RESULT 13

Q16468 PRELIMINARY; PRT; 8 AA.

ID Q16468
AC Q16468;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE DNA for cosmid cc13-1134 PCR primer 1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96435920; PubMed=8838806;
RA James L.A., Ogilvie D.J., Yamakawa K., Nakamura Y., Stirling C.J.,
Rt Anand R.;
RT "Walking, cloning and mapping with YACs in 3q27. Localisation of 5
RT ESTs including 3 members of the Cystatin gene family and
RT identification of CpG islands.";
RL Genomics 32:425-430(1996).
DR EMBL; X88976; CAA61407.1; -.
FT NON_TER 8
FF 8
SQ SEQUENCE 8 AA; 925 MW; FD5411A7376871B6 CRC64;

Query Match 30.6%; Score 11; DB 4; Length 8;
Best Local Similarity 66.7%; Pred No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNN 3
|
|
Db 6 TRN 8

RESULT 14

Q8WNS1 PRELIMINARY; PRT; 8 AA.

ID Q8WNS1
AC Q8WNS1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE X-linked zinc finger protein (fragment).
GN ZFX.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Polounienko A., Blecher S.;
RT "Comparison between intron-exon structures in zFX and zFY genes.";
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF045782; AAL58190.1; -.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 904 MW; DF1DC2C4472AAB1A CRC64;

Query Match 30.6%; Score 11; DB 6; Length 8;
Best Local Similarity 50.0%; Pred No. 6.7e+05;
Matches 2; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 NVLQ 6
|
|
|
Db 5 NVAE 8


```

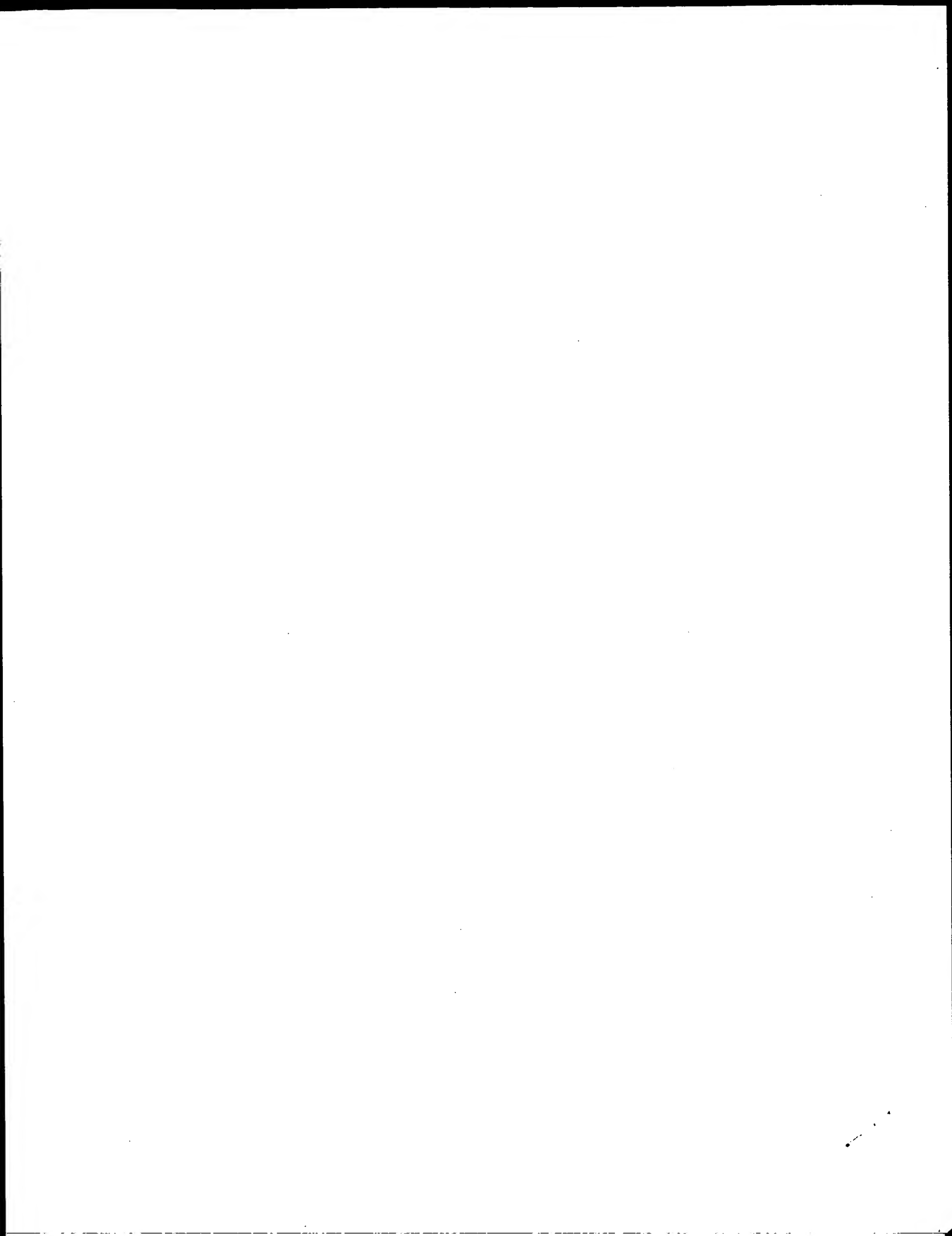
RESULT 15
Q9MSX1
ID Q9MSX1 PRELIMINARY; PRT; 8 AA.
AC Q9MSX1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PsbA (Fragment).
DE PSBA.
GN Jurinea humilis.
OS Jurinea humilis.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Cardueae; Jurinea.
OX NCBI_TaxID=41594;
RN [1]
RP SEQUENCE FROM N.A.
RA O'Hanlon P.C., Briese D.T.;
RT "Evidence for hierarchical and non-hierarchical evolution in the
RT Carduinae thistles";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF129846; AAF78138.1; -.
KW Chloroplast.
FT NON_TER
SQ SEQUENCE 8 AA; 788 MW; 457451B5A76DDB10 CRC64;

Query Match 30.6%; Score 11; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 6.7e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TN 2
   ||
Db 6 TN 7

```

Search completed: February 26, 2003, 15:45:22
 Job time : 93 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:38:41 ; Search time 34 seconds
(without alignments)
31.353 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 86491

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	94.4	8	20 AAW87974	Lectin derived pro
2	34	94.4	8	22 AAG62889	Peptide derived fr
3	20	61.1	6	20 AAY23824	Peptide derived fr
4	20	55.6	7	18 AAW27518	Human interleukin
5	20	55.6	8	15 AAR57168	Peptide corresp. t
6	19	52.8	7	21 AAB23214	Hsp47-binding phag
7	19	52.8	7	22 AAB72780	Antibacterial pept
8	19	52.8	7	22 AAB72821	Antibacterial pept
9	19	52.8	7	22 AAB72823	Antibacterial pept
10	18	50.0	5	15 AAR54171	CHA255 light chain

11	18	50.0	5	21 AAB52719	T21/DP107 peptide
12	18	50.0	5	21 AAB14673	HIV-1 isolate LAI
13	18	50.0	5	22 AAB54937	Anti-HIV peptide D
14	18	50.0	6	20 AAY23409	V beta 6 clone fou
15	18	50.0	6	20 AAW78331	Human fibrinogen g
16	18	50.0	6	21 AAB52720	T21/DP107 peptide
17	18	50.0	6	21 AAB14674	HIV-1 isolate LAI
18	18	50.0	6	22 AAB54938	Anti-HIV peptide D
19	18	50.0	7	21 AAB52721	T21/DP107 peptide
20	18	50.0	7	21 AAB14675	HIV-1 isolate LAI
21	18	50.0	7	21 AAY81478	Bovine membrane-bo
22	18	50.0	7	22 AAB72812	Antibacterial pept
23	18	50.0	7	22 AAB54939	Anti-HIV peptide D
24	18	50.0	8	21 AAB52722	T21/DP107 peptide
25	18	50.0	8	21 AAB14676	HIV-1 isolate LAI
26	18	50.0	8	22 ABP12040	HIV A02 super moti
27	18	50.0	8	22 AAB54940	Anti-HIV peptide D
28	17	47.2	4	21 AAB52718	T20/DP178 peptide
29	17	47.2	4	21 AAB14672	HIV-1 isolate LAI
30	17	47.2	4	22 AAB54936	Anti-HIV peptide D
31	17	47.2	5	15 AAR54176	CHA255 light chain
32	17	47.2	5	15 AAR54167	CHA255 light chain
33	17	47.2	5	15 AAR54173	CHA255 light chain
34	17	47.2	5	15 AAR54175	CHA255 light chain
35	17	47.2	5	19 AAW65349	Peptide #10 useful
36	17	47.2	5	19 AAW65350	Peptide #11 useful
37	17	47.2	5	19 AAW65351	Peptide #12 useful
38	17	47.2	5	21 AAB37736	Fibronectin-derive
39	17	47.2	5	21 AAB37737	Fibronectin-derive
40	17	47.2	5	21 AAB37738	Fibronectin-derive
41	17	47.2	5	21 AAY80733	Fibroblast invasio
42	17	47.2	5	21 AAY80734	Fibroblast invasio
43	17	47.2	5	21 AAY80735	Fibroblast invasio
44	17	47.2	5	22 AAB68154	Meat protein deriv
45	17	47.2	5	22 AAB37554	CD8 activation com

ALIGNMENTS

RESULT 1

AAW87974

ID AAW87974 standard; Peptide; 8 AA.

XX AAW87974;

DT 13-APR-1999 (first entry)

DE Lectin derived progenitor cell preservation factor derived peptide.

KW Lectin derived progenitor cell preservation factor; progenitor cell;
KW haematopoietic cell; cultured cell preservation; anticancer therapy;
KW myeloablative therapy; sickle-cell anaemia; ablative therapy protection;
KW FLK2/FLT3 receptor; ss.

XX Dolichos lab lab.

PN WO9859038-A1.

PD 30-DEC-1998.

XX 23-JUN-1998; 98WO-US13046.

XX 24-JUN-1997; 97US-0881189.

PA (IMCL-) IMCLONE SYSTEMS INC.

PA (REGC) UNIV CALIFORNIA.

PI Chrispeels MJ, Colucci MG, Moore JG;

DR WPI; 1999-081274/07.

PT New nucleic acid encoding plant lectin that preserves progenitor

PT cells - particularly haematopoietic progenitors, useful for bone
PT marrow reconstitution after ablative therapy, and to increase DNA
PT transfer in gene therapy
XX
PS Claim 1; Page 46; 72pp; English.
XX
CC The present sequence represents a peptide of lectin derived progenitor
CC a cell preservation factor. The protein is used to preserve unipotent,
CC pluripotent or totipotent progenitor cells, especially haematopoietic
CC cells, and also progenitors from nerve, muscle, skin, gut, bone,
CC kidney, liver, pancreas or thymus. Specific applications are
CC preservation of cultured cells intended for administration after
CC (anticancer) myeloablative therapy (bone marrow or whole-body irradiation
CC or chemotherapy) to reconstitute the haematopoietic system; enrichment
CC of progenitor cells (e.g. during ex vivo purging of malignant cells);
CC treatment of tissues containing haematopoietic progenitors for subsequent
CC transplant to improve haematopoietic competence; improving transfer of
CC exogenous DNA to progenitor cells (in gene therapy of various
CC haematological disorders, e.g. sickle-cell anaemia); and protection
CC against ablative therapy (to eliminate proliferating cells specifically),
CC followed by re-establishment of differentiation and proliferation of
CC preserved progenitors. The protein, when linked to magnetic beads, may
CC also be used to isolate cells that express the FLK2/FLT3 receptor.
XX
SQ Sequence 8 AA;
Query Match 94.4%; Score 34; DB 20; Length 8;
Best Local Similarity 87.5%; Pred. No. 7.8e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TNNVLQXT 8
| | | | | | | |
Db 1 TNNVLQVT 8

RESULT 2
AAG62889
ID AAG62889 standard; peptide; 8 AA.
XX
AC AAG62889;
XX
DT 17-SEP-2001 (first entry)
XX
DE Peptide derived from a hyacinth bean FRIL polypeptide.
XX
KW FRIL; FLK2/Flt3 tyrosine kinase receptor-interacting lectin;
KW progenitor cell preservation factor; radiotherapy; chemotherapy;
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
KW severe combined immunodeficiency; aplastic anemia; tissue repair.
XX
OS Dolichos lab lab.
XX
FH Key Location/Qualifiers
FT Misc-difference 7
FT /note= "not specified"
XX
PN WO200149851-A1.
XX
PD 12-JUL-2001.
XX
PF 30-DEC-1999; 99WO-US31307.
XX
PR 30-DEC-1999; 99WO-US31307.
XX
PA (PHYL-) PHYLOGIX LLC.
XX
PI Colucci MG, Chrispeels MJ, Moore JG;
XX
DR WPI; 2001-441882/47.
XX
PT Legume progenitor cell preservation factors for in vivo or ex vivo
PT preservation of haematopoietic progenitor cells and as therapeutics for
PT alleviating/reducing progenitor cell-depleting activity of cancer

therapeutics -
Disclosure; Page 19; 173pp; English.
XX
CC The present sequence is derived from a FRIL (FLK2/Flt3 tyrosine kinase
CC receptor-interacting lectin) polypeptide. The specification describes a
CC composition of one or more members of FRIL family of progenitor cell
CC preservation factors. The composition is useful for alleviating or
CC reducing the hematopoietic progenitor cell-depleting activity of a
CC therapeutic treatment, including radiotherapeutic and/or
CC chemotherapeutic treatments. Administration of FRIL compositions to a
CC patient prior to treatment of the patient with a therapeutic treatment
CC having a hematopoietic progenitor cell-depleting activity alleviates or
CC reduces the hematopoietic progenitor cell-depleting activity of the
CC therapeutic treatment in the patient. FRIL family members are useful for
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
CC stem cells. The composition is administered to reduce progenitor cell
CC depleting effects of chemotherapeutics, so that the patient can receive
CC a higher dose of the chemotherapeutic and preferably recover from cancer.
CC It is also administered to patients having, or predisposed to developing
CC a condition where the patients hematopoietic progenitor cells are
CC depleted, such as severe combined immunodeficiency or aplastic anemia.
CC The isolated mesenchymal cells are useful for tissue repair.
XX
SQ Sequence 8 AA;
Query Match 94.4%; Score 34; DB 22; Length 8;
Best Local Similarity 100.0%; Pred. No. 7.9e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNNVLQXT 8
| | | | | | | |
Db 1 TNNVLQVT 8

RESULT 3
AAY23824
ID AAY23824 standard; Peptide; 6 AA.
XX
AC AAY23824;
XX
DT 14-SEP-1999 (first entry)
XX
DE Peptide derived from gyrase protein and used to design primers.
XX
KW Identification; detection; microbe; gyrase gene; gyrase protein.
XX
OS Unidentified.
XX
PN JF11169175-A.
XX
PD 29-JUN-1999.
XX
PF 12-DEC-1997; 97JP-0343316.
XX
PR 12-DEC-1997; 97JP-0343316.
XX
PA (MARI-) MARINE BIOTECHNOLOGY INST CO LTD.
XX
PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
XX
DR WPI; 1999-422615/36.
XX
PT Identification and detection of a microbe - by detection of a
PT gyrase gene
XX
PS Example 2; Page 36; 42pp; Japanese.
XX
CC The specification describes a method for the identification or
CC detection of a microbe, using the gyrase gene as the index. The
CC method involves the use of PCR primers to amplify DNA from the
CC microbe, which is then identified or detected depending on its
CC base sequence. The method can be used to classify and identify
CC an unidentified microbe strain rapidly and with high precision.

CC The present sequence represents a peptide derived from gyrase
 CC protein and used to design primers for use in the method of the
 CC invention.
 XX
 XX
 SQ Sequence 6 AA;

Query Match 61.1%; Score 22; DB 20; Length 6;
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 |||:
 Db 1 TNNIPQ 6

RESULT 4
 AAW27518
 ID AAW27518 standard; peptide; 7 AA.

XX AC AAW27518;
 XX 09-JAN-1998 (first entry)
 XX Human interleukin 9 receptor antigenic peptide.

XX Antagonist; human; interleukin 9; asthma associated factor 1; IL-9;
 KW atopic allergy; asthma; bronchial hypersensitivity; BHR; eczema;
 KW rhinitis; urticaria; allergic inflammation; bowel; alleviation;
 KW treatment; antigen; antibody; receptor.

XX Homo sapiens.
 XX WO9708321-A1.
 XX 06-MAR-1997.

XX 23-AUG-1996; 96WO-US12757.
 XX 06-AUG-1996; 96US-0023800.
 XX 24-AUG-1995; 95US-0002765.

XX (MAGA-) MAGAININ PHARM INC.

XX Lee MW, Levitt RC, Nicholas N, Prasad KU;

XX WPI; 1997-179278/16.

XX Human interleukin-9 variant with Met at position 117 - useful for
 XX treating atopic allergy, esp. asthma

XX Claim 105; Page 105; 142pp; English.

XX The present peptide is an antigen derived from human interleukin 9
 CC (hIL-9) receptor, also known as asthma associated factor 1, which
 CC can be used to raise antibodies (Ab) that block the binding of
 CC hIL-9 to its receptor. hIL-9 plays a role in atopic allergy,
 CC asthma and related disorders, e.g. bronchial hypersensitivity
 CC (BHR), rhinitis, urticaria, allergic inflammatory disorders of the
 CC bowel and various forms of eczema. Compounds which down regulate
 CC the function of IL-9, e.g. antisense DNA, Ab, truncated (soluble)
 CC IL-9 and its splice variants and other IL-9 or IL-9 receptor
 CC antagonists, are useful for alleviating such diseases. Anti-IL-9
 CC Ab are also useful for quantifying IL-9 in a sample.

XX SQ Sequence 7 AA;

Query Match 55.6%; Score 20; DB 18; Length 7;
 Best Local Similarity 75.0%; Pred. No. 7.8e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
 |||:
 Db 4 TNNI 7

RESULT 5
 AAR57168
 ID AAR57168 standard; Protein; 8 AA.

XX AC AAR57168;

XX 28-FEB-1995 (first entry)

XX Peptide corresp. to E.coli CS31A subunit residues 88-96.

XX Capsule protein; CS31A; surface antigen; immunogen; vaccine;
 KW insertion; modification; mutation; epitope; poliovirus;
 KW porcine transmissible gastroenteritis virus; bovine rotavirus;
 KW foot and mouth disease virus; clpG gene.

XX Escherichia coli (strain 31A).

XX FR2699538-A.

XX 24-JUN-1994.

XX 22-DEC-1992; 92FR-0015464.

XX 22-DEC-1992; 92FR-0015464.

XX (INRG) INRA INST NAT RECH AGRONOMIQUE.

XX Bousquet F, Der Vartanian M, Girardeau J, Martin C;

XX Mechlin M;

XX WPI; 1994-227158/28.

XX Sub-unit of capsular protein CS31A contg heterologous sequences -
 PT useful for representation of heterologous epitopes in vaccines,
 PT for antibody prodn and in immunological tests

XX Disclosure; Page 36; 128pp; French.

XX The CS31A subunit of the E.coli capsular protein can accept large
 CC inserts (up to 100 amino acids) without inhibition of biosynthesis.
 CC The insertions are esp. epitope sequences from viruses and the
 CC modified subunits, or microorganisms expressing them, are useful as
 CC immunogens. Pref. the subunit has at least one insertion and/or
 CC substitution by a heterologous peptide in one of the regions -13 to
 CC +8, 10-58, 123-164 or 183-257. Esp. the modifications are made in
 CC the regions corresp. to natural epitopes of CS31A as these are the
 CC most accessible sites on the folded protein. The epitopes were
 CC identified and localised by assaying antipeptide capture.
 CC This peptide was shown to correspond to an accessible region.

XX SQ Sequence 8 AA;

Query Match 55.6%; Score 20; DB 15; Length 8;
 Best Local Similarity 57.1%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 TNNVLQXT. 8
 |||:
 Db 2 NGVLQSS 8

RESULT 6
 AAB23214

ID AAB23214 standard; peptide; 7 AA.

XX AC AAB23214;

XX 29-JAN-2001 (first entry)

XX Hsp47-binding phage display heptapeptide, SEQ ID NO:36.
 DE
 XX

KW Hsp47-binding heptapeptide; phage display library;
 KW Hsp47 external domain; carcinoma; cancer; targeting molecule;
 KW therapy; diagnosis; detection; imaging; drug delivery; invasion;
 KW migration; metastasis; modulation; tumour; skin; basal cells; colon;
 KW large intestine; lung; breast; bladder; oral cancer;
 KW head and neck cancer; larynx; nasopharynx; adrenal cortex;
 KW apocrine gland; kidney; liver; pancreas; prostate.
 XX
 OS Synthetic.
 XX
 XX WO2000054805-A1.
 PN
 XX
 XX 21-SEP-2000.
 PD
 XX
 PF 15-MAR-2000; 2000WO-US06588.
 XX
 XX 15-MAR-1999; 99US-0124481.
 PR
 XX
 PA (UYNA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Sauk JJ;
 XX
 DR WPI; 2000-655997/63.
 XX
 PT Treating, diagnosing or modulating a carcinoma cell, which expresses
 PT Heat shock protein 47 on its surface, involves administering an agent
 PT comprising targeting moiety which binds to Hsp47 external domain -
 PS
 XX Example 5; Page 9; 87pp; English.
 XX
 CC The invention relates to methods of treating and diagnosing carcinomas
 CC in which heat shock protein 47 (Hsp47) is expressed on the surface
 CC of the carcinoma cells, involving administering an agent comprising a
 CC targeting moiety which specifically binds to the external domain of
 CC Hsp47. The invention also relates to peptides (AAB23181-B23203) which
 CC specifically bind to external domains of such surface-localised Hsp47
 CC molecules and have sequences encompassed by the generic sequences
 CC XHVVHXXHXXHXXHXX or HyXXXHXXHXXHXX where X, independently, can be
 CC any amino acid and Hy, independently, can be any hydrophobic amino acid.
 CC The invention also encompasses methods of screening for agents which
 CC bind Hsp47 external domains. Hsp47-binding agents can be used to
 CC treat Hsp47-expressing carcinomas, and for modulating the activity of a
 CC tumour cell with respect to invasion, migration, motility or metastasis,
 CC or to its interaction with the extracellular matrix. The targeting
 CC moiety (such as an Hsp47-binding peptide, a toxin or an antibody) may be
 CC coupled with a therapeutic moiety (such as a cytotoxic agent or a
 CC therapeutic gene) for cancer treatment, or with a detectable moiety for
 CC imaging. Carcinomas which may be treated or diagnosed according to
 CC methods of the invention include those of the skin, basal cells, large
 CC intestine, lung, colon, breast, bladder, oral, head and neck, larynx,
 CC nasopharynx, adrenal cortex, apocrine glands, kidney, liver, pancreas, or
 CC prostate. Targeting carcinoma cells with Hsp47-binding agents results in
 CC efficient delivery of therapeutic agents, reduced doses, reduced side
 CC effects and sensitive detection or imaging of carcinoma cells. Sequences
 CC AAB23204-B23240 represent phage display library heptapeptides identified
 CC as being able to bind Hsp47 in an exemplification of the invention.
 XX
 XX Sequence 7 AA;
 SQ

Query Match 52.8%; Score 19; DB 21; Length 7;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVL 5
 DB 3 TNNVL 7

RESULT 7
 AAB72780
 ID AAB72780 standard; Peptide; 7 AA.
 XX
 AC AAB72780;

XX 09-MAY-2001 (first entry)
 DT Antibacterial peptide SEQ ID NO: 1.
 DE
 XX
 XX Antibacterial compound; Gram-negative bacterium; pilus; chaperone;
 KW biofilm; disease treatment; bacterial infection.
 KW
 XX Synthetic.
 OS
 XX WO200110386-A2.
 PN
 XX
 XX 15-FEB-2001.
 PD
 XX
 PF 11-AUG-2000; 2000WO-US22087.
 XX
 XX 11-AUG-1999; 99US-0148280.
 PR
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 XX Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;
 PI WPI; 2001-226496/23.
 XX
 DR An isolated compound for inhibiting pilus assembly -
 XX
 PT Claim 12; Page 130; 144pp; English.
 XX
 CC The present invention provides antibacterial compounds which are able to
 CC interfere with Gram-negative bacteria pilus formation and assembly, and
 CC pilus interaction with chaperone proteins. These are useful in the
 CC treatment of bacterial infection, and in the prevention of biofilm
 CC formation. They are particularly useful against *Escherichia coli*,
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and
 CC *Klebsiella pneumoniae*.
 XX
 SQ Sequence 7 AA;
 Query Match 52.8%; Score 19; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
 DB 1 NVLQ 4

RESULT 8
 AAB72821
 ID AAB72821 standard; Peptide; 7 AA.
 XX
 XX AAB72821;
 AC
 XX 09-MAY-2001 (first entry)
 DT
 XX
 XX Antibacterial peptide SEQ ID NO: 42.
 DE
 XX Antibacterial compound; Gram-negative bacterium; pilus; chaperone;
 KW biofilm; disease treatment; bacterial infection.
 KW
 XX Synthetic.
 OS
 XX WO200110386-A2.
 PN
 XX
 XX 15-FEB-2001.
 PD
 XX
 PF 11-AUG-2000; 2000WO-US22087.
 XX
 XX 11-AUG-1999; 99US-0148280.
 PR
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX

PI Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;
 XX WPI; 2001-226496/23.
 XX
 XX An isolated compound for inhibiting pilus assembly -
 XX
 XX Claim 12; Page 140; 144pp; English.

XX The present invention provides antibacterial compounds which are able to
 CC interfere with Gram-negative bacteria pilus formation and assembly, and
 CC pilus interaction with chaperone proteins. These are useful in the
 CC treatment of bacterial infection, and in the prevention of biofilm
 CC formation. They are particularly useful against *Escherichia coli*,
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and
 CC *Klebsiella pneumoniae*.

XX Sequence 7 AA;

Query Match 52.8%; Score 19; DB 22; Length 7;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
 DB 1 NVLQ 4

RESULT 9
 AAB72823
 ID AAB72823 standard; Peptide; 7 AA.

XX AC AAB72823;
 XX 09-MAY-2001 (first entry)

XX Antibacterial peptide SEQ ID NO: 44.

XX Antibacterial compound; Gram-negative bacterium; pilus; chaperone;
 KW biofilm; disease treatment; bacterial infection.

XX Synthetic.

XX WO200110386-A2.

XX 15-FEB-2001.

XX 11-AUG-2000; 2000WO-US22087.

XX 11-AUG-1999; 99US-0148280.

XX (UNIW) UNIV WASHINGTON.

XX Hultgren SJ, Sauer FG, Waksman G, Fuetterer K;
 XX WPI; 2001-226496/23.

XX An isolated compound for inhibiting pilus assembly -

XX Claim 12; Page 141; 144pp; English.

XX The present invention provides antibacterial compounds which are able to
 CC interfere with Gram-negative bacteria pilus formation and assembly, and
 CC pilus interaction with chaperone proteins. These are useful in the
 CC treatment of bacterial infection, and in the prevention of biofilm
 CC formation. They are particularly useful against *Escherichia coli*,
 CC *Haemophilus influenzae*, *Salmonella enteritidis*, *S. typhimurium*,
 CC *Bordetella pertussis*, *Yersinia enterocolitica*, *Helicobacter pylori* and
 CC *Klebsiella pneumoniae*.

XX Sequence 7 AA;

Query Match 52.8%; Score 19; DB 22; Length 7;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
 DB 1 NVLQ 4

RESULT 10

AAR54171
 ID AAR54171 standard; Protein; 5 AA.

XX AC AAR54171;

XX 22-FEB-1995 (first entry)

XX CHA255 light chain CDR2 clone 4.

XX Polymerase chain reaction; primer; PCR; amplify; heavy; light;
 KW chain; complementarity determining region; CDR; variable; constant;
 KW region; monoclonal antibody; MAB; binding affinity; EDTA; DOTA;
 KW tumour; cancer; colorectal; breast; metal chelate; haptene.

XX Synthetic.

XX AU9350602-A.

XX 26-MAY-1994.

XX 10-NOV-1993; 93AU-0050602.

XX 12-NOV-1992; 92US-0975230.

XX (HYBR-) HYBRITECH INC.

XX Ahrweiler PM, Moore MD;

XX WPI; 1994-209063/26.

XX N-PSDB; AAQ68773.

XX Polypeptide used in imaging and treatment of carcinomas and
 PT tumours - comprising subunit antibody CDR having binding affinity
 PT for metal chelate of EDTA or DOTA or analogues

XX Claim 25; Fig 3B; 61pp; English.

XX The sequences given in AAR54167-76 represent the wild type and
 CC mutagenised versions of the complementarity determining region 2 (CDR2)
 CC of the antibody designated CHA255 light chain. CHA255 is a murine
 CC monoclonal antibody (MAB) which is capable of binding complexes.
 CC Mutagenesis of these CDRs, causes the production of polypeptides with a
 CC particularly high binding affinity for EDTA or DOTA metal complexes.
 CC CDR1 and -3 of the heavy chain, and CDR2 and -3 of the light chain were
 CC targeted for mutagenesis. Five residues of both CDR1 and -3 of the
 CC CHA255 heavy chain, five of seven residues of light chain CDR and six of
 CC nine light chain CDR3 residues were specifically targeted for
 CC codon-based mutagenesis. The mutagenised MAB's can be used in
 CC compositions for in vivo imaging of malignant tissues or tumours. They
 CC are also useful for the treatment of malignant tissues or tumours eg.
 CC colorectal or breast cancer. Both methods involve the use of
 CC radionuclides which bind to metal chelates or haptens which are
 CC specifically delivered to the target site by a targeting molecule. CDR
 CC derived peptides may be used to construct bi-functional antibodies
 CC having dual specificities, or as donor or recipients of CDR sequences.

XX Sequence 5 AA;

Query Match 50.0%; Score 18; DB 15; Length 5;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNVNL 5
 IIII

Db 1 TNNRL 5

RESULT 11
AAB52719
ID AAB52719 standard; Peptide; 5 AA.
XX AC AAB52719;
XX DT 23-FEB-2001 (first entry)
XX DE T21/DP107 peptide fragment #1.
XX KW Antinflammatory; T21/DP107; gp41 ectodomain; HIV-1 fusion;
KW formyl peptide receptor family; FPR; inflammatory response up-regulation;
KW chemoattractant.
XX OS Human immunodeficiency virus type 1.
XX PN WO200066622-A1.
XX PD 09-NOV-2000.
XX PF 05-MAY-2000; 2000WO-US12371.
XX PR 05-MAY-1999; 99US-0132686.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Wang JM, Oppenheim JJ, Su S, Gong W, Gao J, Murphy PM;
XX WPI; 2000-656493/63.
XX DR Administration of peptide agents with a sequence corresponding to a
PT partial-length T20/DP178 or T21/DP107 formyl peptide receptor
PT antagonist is used to modulate inflammation -
XX Claim 12; Page 27; 148pp; English.
XX The present sequence is a peptide fragment of T21/DP107. T21/DP107 is a
CC helical segment of the ectodomain of HIV-1 protein gp41. T21/DP107 is
CC located in the amino terminus of the gp41 ectodomain. gp41 plays a
CC critical role in the fusion of HIV-1 and host cell membranes. T21/DP107
CC interacts with members of the formyl peptide receptor (FPR) family and
CC thereby up-regulates an inflammatory response, and acts as a potent
CC chemoattractant and activator of human peripheral blood phagocytes
CC (but not T cells). The present peptide can be used to modulate an
CC inflammatory response in a subject.
XX Sequence 5 AA;
SQ Query Match 50.0%; Score 18; DB 21; Length 5;
Best Local Similarity 60.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 NNVLQ 6
Db 1 NNLRL 5
RESULT 12
AAB14673
ID AAB14673 standard; peptide; 5 AA.
XX AC AAB14673;
XX DT 24-NOV-2000 (first entry)
XX DE HIV-1 isolate LAI gp41 N-helical domain peptide P-17, fragment #2.
XX KW HIV-1; gp41 N-helical domain; gp41 heptad repeat region;
KW coiled coil; core 6-helix bundle; viral entry inhibition; immunogenic;
KW antibody; humoral response; broad spectrum vaccine; anti-HIV;

KW envelope glycoprotein; prophylaxis; therapy; group M; subtype B;
KW isolate LAI.
OS Human immunodeficiency virus type 1.
XX WO200040616-A1.
XX PD 13-JUL-2000.
XX PF 10-JAN-2000; 2000WO-US00456.
XX PR 08-JAN-1999; 99US-0115404.
XX PF 07-JAN-2000; 2000US-0480336.
XX PA (WILD/) WILD C T.
XX PA (WEIS/) WEISS C D.
XX PI Wild CT, Weiss CD;
XX DR WPI; 2000-465959/40.
XX Raising neutralizing antibody response to human immunodeficiency virus,
PT comprises administering a polypeptide capable of forming a stable
PT coiled-coil solution structure -
XX Disclosure; Page 35; 97pp; English.
XX Sequences AAB14672-B14739 represent peptides derived from the
CC N-helical domain of the gp41 envelope glycoprotein from HIV-1
CC group M, subtype B, isolate LAI. The invention relates to raising
CC a neutralising antibody response to a broad spectrum of HIV (human
CC immunodeficiency virus) strains and isolates, comprising the
CC administration of a peptide which corresponds to or mimics highly
CC conserved portions of gp41 which are important in mediating the
CC process of viral entry into host cells. Such peptides can correspond
CC to or mimic the coiled coil solution structure of the N-helical domain
CC (the heptad repeat region), or can correspond or mimic the C-helical
CC domain (the transmembrane-proximal amphipathic alpha-helical segment),
CC or the gp41 core 6-helix bundle, which is formed by the interaction
CC of the N- and C-helical domains of three gp41 proteins. The peptides
CC can be administered either singly or as a combination (particularly
CC a combination of N-helical and C-helical peptides), and can be
CC multimerised. For example, N- and C-helical domain peptides can be
CC alternately linked together to form a peptide which mimics the core
CC 6-helix bundle. Administration of the peptide(s) generates a humoral
CC response, with the production of antibodies against gp41 structures
CC involved in viral entry. As these portions of gp41 are well conserved,
CC such antibodies may be effective against a broad range of HIV strains
CC and isolates. The peptide compositions may be administered as a
CC prophylactic or therapeutic vaccine to generate antibodies which reduce
CC or inhibit the ability of HIV to infect uninfected cells. A composition
CC comprising polyclonal or monoclonal antibodies can be administered to
CC reduce HIV infection of uninfected cells. Antibodies raised against
CC entry-relevant gp41 structures may also be used therapeutically and as
CC tools to further elucidate the mechanism of HIV cell entry.
XX Sequence 5 AA;
SQ Query Match 50.0%; Score 18; DB 21; Length 5;
Best Local Similarity 60.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 NNVLQ 6
Db 1 NNLRL 5
RESULT 13
AAB54937
ID AAB54937 standard; Peptide; 5 AA.
XX AC AAB54937;
XX XX

DT 05-MAR-2001 (first entry)
 XX
 DE Anti-HIV peptide DPl07 carboxy truncation peptide #2.
 XX
 KW Long lasting fusion peptide inhibitor; viral infection; antiviral;
 KW antifusogenic; mobile blood component; measles virus; MeV; SIV;
 KW smian immunodeficiency virus; human parainfluenza virus; HPIV; RSV;
 KW human respiratory syncytial virus; human immunodeficiency virus; HIV.
 XX
 OS Human immunodeficiency virus type 1.
 XX
 XX WO200069902-A1.
 PN
 XX 23-NOV-2000.
 PD
 XX 17-MAY-2000; 2000WO-US13651.
 XX
 PF 17-MAY-1999; 99US-0134406.
 XX
 PR 10-SEP-1999; 99US-0153406.
 PR
 XX (CONJ-) CONJUCHEM INC.
 PA
 PA Bridon DP, Dufresne RP, Boudjellab N, Robitaille M, Milner PG;
 XX
 PI WPI; 2001-007496/01.
 XX
 DR A modified peptide and a reactive group which is reactive with amino
 XX groups, hydroxyl groups, or thiol groups on blood components to form
 PT stable covalent bonds useful for treatment of viral infections, e.g.,
 PT human immunodeficiency virus -
 PT
 XX Disclosure; Page 136; 21lpp; English.
 PS
 XX The present invention describes a modified anti-viral peptide (I)
 CC comprising a peptide that exhibits anti-viral activity and a reactive
 CC group which is reactive with amino groups, hydroxyl groups, or thiol
 CC groups on blood components to form stable covalent bonds. (I) has
 CC anti-viral and anti-fusogenic activities. (I) inhibits viral infection
 CC of cells by inhibiting cell-cell fusion or free virus infection or to
 CC reduce the level of membrane fusion events between two or more entities,
 CC e.g., virus-cell or cell-cell, relative to the level of membrane fusion
 CC that occurs in the absence of the peptide. (I) is useful in the
 CC treatment of patients who are suffering from viral infection, e.g. HIV,
 CC RSV, HPIV, MeV, and SIV. (I) may be administered prophylactically to
 CC previously uninfected individuals. This is useful in cases where an
 CC individual has been subjected to a high risk of exposure to a virus.
 CC By bonding of long-lived components of the blood, such as immunoglobulin,
 CC serum albumin, red blood cells and platelets the activity is extended
 CC for days to weeks. This is due to improved stability in vivo and a
 CC reduced susceptibility to peptidase or protease degradation. This
 CC minimises the need for more frequent, or even continual, administration
 CC of the peptides. AAB54784 to AAB55431 represent peptides used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 5 AA;
 Query Match 50.0%; Score 18; DB 22; Length 5;
 Best Local Similarity 60.0%; Pred. No. 7.Be+05;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 2 NNVLQ 6
 DB 1 NNLLR 5
 RESULT 14
 AAY23409
 ID AAY23409 standard; Peptide; 6 AA.
 XX
 AC AAY23409;
 XX
 DT 02-SEP-1999 (first entry)
 XX

DE V beta 6 clone found in MS patients after vaccination with TCR.
 XX
 KW Vaccine; T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;
 KW V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;
 KW multiple sclerosis.
 XX
 KW Synthetic.
 OS
 OS Homo sapiens.
 XX
 PN WO9927957-A1.
 XX
 XX 10-JUN-1999.
 PD
 XX 03-DEC-1997; 97WO-US23147.
 XX
 PF 03-DEC-1997; 97WO-US23147.
 XX
 PR (IMU-) IMMUNE RESPONSE CORP.
 XX
 PA (KIMM-) KIMMEL CANCER CENT SIDNEY.
 PA
 XX Brostoff SW, Carlo DJ, Gold DP, Smith LR, Wilson DB;
 PI WPI; 1999-404801/34.
 XX
 DR T0 cell receptor peptide-derived vaccines
 XX
 PT Example 11; Page 85; 104pp; English.
 PS
 XX The specification describes vaccines which comprise immunologically
 CC effective amounts of T cell receptor (TCR) peptides. The TCRs are
 CC present on the surface of T cells. The TCRs are chosen from V beta
 CC 6.2/3, V beta 6/5, V beta 6.7, V beta 2, V beta 5/1, V beta 7 or V beta
 CC 13. The V beta TCR peptide-based vaccines are useful for prevention or
 CC treatment of multiple sclerosis (MS). The presence of V beta 6.7 appears
 CC to be particularly associated with multiple sclerosis and can be used
 CC to determine an individual's susceptibility to multiple sclerosis.
 CC Vaccinating, rather than passively administering heterologous
 CC antibodies, allows the host's own immune system to mobilize and suppress
 CC auto aggressive T cells. Therefore, the suppression is persistent and
 CC may involve any and all immunological mechanisms in effecting that
 CC suppression. Such a multi-faceted response is more effective than
 CC the uni-dimensional suppression achieved by passive administration of
 CC monoclonal antibodies or extant-derived regulatory T cell clones.
 CC AAY23387-Y23480 represent peptides derived from TCR V beta 6 clones
 CC found in the cerebrospinal fluid (CSF) of MS patients, after vaccination
 CC with V beta 6.
 XX
 SQ Sequence 6 AA;
 Query Match 50.0%; Score 18; DB 20; Length 6;
 Best Local Similarity 60.0%; Pred. No. 7.8e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 NNVLQ 6
 DB 1 NNALE 5
 RESULT 15
 AAW78331
 ID AAW78331 standard; peptide; 6 AA.
 XX
 AC AAW78331;
 XX
 XX 27-APR-1999 (first entry)
 DT
 XX Human fibrinogen gamma A/B chain protein fragment.
 DE
 XX Expression; epitope; recombinant vector; structural gene; bacteriophage;
 KW identification.
 KW
 XX Homo sapiens.
 OS
 XX

```

PN US5866363-A.
XX
PD 02-FEB-1999.
XX
PF 28-FEB-1991; 91US-0662764.
XX
PR 28-FEB-1991; 91US-0662764.
PR 28-AUG-1985; 85US-0770390.
PR 26-MAY-1988; 88US-0201358.
XX
PA (PIEC/) PIECZENIK G.
XX
PI Pieczenik G;
XX
DR WPI: 1999-141923/12.
DR N-PSDB; AAX16850.
XX
PT Population of recombinant vectors, oligonucleotides and random
PT peptide sequences - useful for the identification and
PT characterisation of peptide epitopes
XX
PS Example 4; Column 21; 26pp; English.
XX
CC The invention relates to the expression of epitope peptides from
CC recombinant vectors which contain a recombinant structural gene into
CC which a random oligonucleotide of sequence GATCCTTN(15)AA has been
CC inserted. Sequences X16832-X16845 represent examples of the inserted
CC oligonucleotides. Several of the peptides encoded by the oligonucleotides
CC have been shown to have similarity to proteins in the GenBank sequence
CC database. This sequence is a fragment of the human fibrinogen gamma A/B
CC chain precursor protein and is example of a homologous peptide sequence.
CC The recombinant peptides are then expressed on the surface of an organism
CC e.g. a bacteriophage, for ease of identification. The invention may be
CC used for the identification and characterisation of peptide epitopes.
XX
SQ Sequence 6 AA;

Query Match 50.0%; Score 18; DB 20; Length 6;
Best Local Similarity 75.0%; Pred. No. 7.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
Db 1 NILQ 4

Search completed: February 26, 2003, 15:43:26
Job time : 35 secs

```

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:45:27 ; Search time 12 Seconds
(without alignments)
25.148 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 16185

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	61.1	6	9	US-09-823-823-61
2	22	61.1	6	10	US-09-823-829-61
3	20	55.6	7	9	US-09-823-823-80
4	20	55.6	7	10	US-09-848-585-22
5	20	55.6	7	10	US-09-823-829-80
6	17	47.2	6	9	US-09-727-963A-56
7	17	47.2	6	10	US-09-990-762-55
8	17	47.2	7	9	US-09-956-086-18
9	17	47.2	7	9	US-09-956-087-18
10	17	47.2	7	9	US-09-995-973-20
11	17	47.2	7	9	US-09-995-973-24
12	17	47.2	7	9	US-09-995-973-27
13	17	47.2	7	10	US-09-865-483-10
14	17	47.2	8	10	US-09-924-703-12
15	16	44.4	6	10	US-09-990-762-54
16	16	44.4	6	10	US-09-911-838-142
17	16	44.4	6	10	US-09-911-838-144
18	16	44.4	7	10	US-09-911-838-141
19	16	44.4	7	10	US-09-911-838-143

20	16	44.4	7	10	US-09-911-838-145
21	16	44.4	8	10	US-09-765-086-157
22	16	44.4	8	10	US-09-950-313-43
23	15	41.7	4	10	US-09-071-838-231
24	15	41.7	7	8	US-08-967-573A-22
25	15	41.7	7	9	US-10-080-100-59
26	15	41.7	7	9	US-10-080-100-100
27	15	41.7	7	10	US-09-989-789-3047
28	15	41.7	7	10	US-09-989-789-3049
29	15	41.7	7	10	US-09-870-379-8
30	15	41.7	8	9	US-10-080-100-45
31	14	38.9	4	10	US-09-220-920-71
32	14	38.9	5	9	US-09-788-006-164
33	14	38.9	6	10	US-09-947-387-121
34	14	38.9	6	10	US-09-911-838-140
35	14	38.9	7	10	US-09-989-789-271
36	14	38.9	7	10	US-09-989-789-331
37	14	38.9	7	10	US-09-989-789-370
38	14	38.9	7	10	US-09-989-789-1013
39	14	38.9	7	10	US-09-989-789-1169
40	14	38.9	7	10	US-09-900-147-7
41	14	38.9	7	10	US-09-947-387-32
42	14	38.9	7	10	US-09-911-838-139
43	14	38.9	8	9	US-09-826-290-53
44	14	38.9	8	9	US-09-826-290-135
45	14	38.9	8	9	US-09-826-290-269

ALIGNMENTS

RESULT 1

US-09-823-823-61
; Sequence 61, Application US/09823823
; Patent No. US20020171092A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Satoshi
; APPLICANT: Kasai, Hiroaki
; APPLICANT: Nakamura, Shoko
; APPLICANT: Suzuki, Makoto
; APPLICANT: Hamoda, Tohru
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING
; FILE REFERENCE: 12817-004001
; CURRENT APPLICATION NUMBER: US/09/823,823
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/208,688
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: JP 97/343316
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patentin version 2.0
; SEQ ID NO 61
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated protein
US-09-823-823-61

Query Match 61.1%; Score 22; DB 9; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.5e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:|
Db 1 TNNIPQ 6

RESULT 2

US-09-823-829-61
; Sequence 61, Application US/09823829
; Patent No. US20020146697A1

GENERAL INFORMATION:
; APPLICANT: Yamamoto, Satoshi
; APPLICANT: Nakamura, Shoko
; APPLICANT: Suzuki, Makoto
; APPLICANT: Kasai, Hiroaki
; APPLICANT: Hamada, Tohru
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS
; FILE REFERENCE: 12817-004001
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/823,829
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: JP 97/343316
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 61
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated peptide
US-09-823-829-61

Query Match 61.1%; Score 22; DB 10; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.5e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
|||:
Db 1 TNNIPQ 6

RESULT 3

US-09-823-823-80
; Sequence 80, Application US/09823823
; Patent No. US20020171092A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Satoshi
; APPLICANT: Kasai, Hiroaki
; APPLICANT: Nakamura, Shoko
; APPLICANT: Suzuki, Makoto
; APPLICANT: Hamada, Tohru
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING G
; FILE REFERENCE: 12817-004001
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/823,823
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: JP 97/343316
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 80
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Xaa at position 6 = Pro or Asn; Xaa at position 7 = Thr or Gln
US-09-823-823-80

Query Match 55.6%; Score 20; DB 9; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
|||:
Db 2 TNNI 5

RESULT 4

US-09-848-585-22
; Sequence 22, Application US/09848585
; Patent No. US20020146391A1
; GENERAL INFORMATION:
; APPLICANT: LEVITT, Roy Clifford
; APPLICANT: MALOY, W. Lee
; APPLICANT: NICOLAIDES, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating Atopic Aller
; FILE REFERENCE: 36870-5056-12-US
; CURRENT FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/002,765
; PRIOR FILING DATE: 1995-08-24
; PRIOR APPLICATION NUMBER: US 08/697,419
; PRIOR FILING DATE: 1996-08-23
; PRIOR APPLICATION NUMBER: US 08/874,503
; PRIOR FILING DATE: 1997-06-13
; PRIOR APPLICATION NUMBER: US 09/325,571
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Residues 8-14 of Mature hIL-9 Receptor
US-09-848-585-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
|||:
Db 4 TNNI 7

RESULT 5

US-09-823-829-80
; Sequence 80, Application US/09823829
; Patent No. US20020146697A1
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Satoshi
; APPLICANT: Nakamura, Shoko
; APPLICANT: Suzuki, Makoto
; APPLICANT: Kasai, Hiroaki
; APPLICANT: Hamada, Tohru
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS
; FILE REFERENCE: 12817-004001
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 09/823,829
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: JP 97/343316
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn version 2.0
; SEQ ID NO 80
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetically generated peptide
; NAME/KEY: VARIANT
; LOCATION: 6
; OTHER INFORMATION: Xaa = Pro or Asn
; NAME/KEY: VARIANT
; LOCATION: 7
; OTHER INFORMATION: Xaa = Thr or Gln

US-09-823-829-80

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 2 TNNI 5

RESULT 6

US-09-727-963A-56
; Sequence 56, Application US/09727963A
; Patent No. US20020155106A1

; GENERAL INFORMATION:

; APPLICANT: V.I. Technologies, Inc.

; APPLICANT: Hammond, David J.

; TITLE OF INVENTION: METHOD OF IDENTIFYING A LIGAND FOR A TARGET MOLECULE

; FILE REFERENCE: 18242-505

; CURRENT APPLICATION NUMBER: US/09/727,963A

; CURRENT FILING DATE: 2000-12-01

; PRIOR APPLICATION NUMBER: U.S.N. 60/327,557

; PRIOR FILING DATE: 1999-12-02

; NUMBER OF SEQ ID NOS: 99

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 56

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Peptide Ligand
US-09-727-963A-56

Query Match

47.2%; Score 17; DB 9; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3
Db 4 TNN 6

RESULT 7

US-09-990-762-55

; Sequence 55, Application US/09990762

; Patent No. US20020119498A1

; GENERAL INFORMATION:

; APPLICANT: JOUNG, J. KEITH

; APPLICANT: MILLER, JEFFREY

; APPLICANT: PABO, CARL O.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INTERACTION TRAP ASSAYS

; FILE REFERENCE: MTV-030.02 (20021-3002)

; CURRENT APPLICATION NUMBER: US/09/990,762

; CURRENT FILING DATE: 2001-11-14

; PRIOR APPLICATION NUMBER: 09/858,852

; PRIOR FILING DATE: 2001-05-16

; PRIOR APPLICATION NUMBER: 60/204,509

; PRIOR FILING DATE: 2000-05-16

; NUMBER OF SEQ ID NOS: 97

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 55

; LENGTH: 6

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Zinc finger

; OTHER INFORMATION: recognition sequence

US-09-990-762-55

Query Match

47.2%; Score 17; DB 10; Length 6;

Best Local Similarity 75.0%; Pred. No. 1.5e+05;

Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 3 SNNV 6

RESULT 8

US-09-956-086-18

; Sequence 18, Application US/09956086

; Patent No. US20020155498A1

; GENERAL INFORMATION:

; APPLICANT: FILPULA, DAVID

; APPLICANT: WANG, MAOLIANG

; APPLICANT: SHORR, ROBERT

; APPLICANT: WHITLOW, MARC

; APPLICANT: LEE, LIHSYNG S.

; TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS

; CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.

; STREET: 1100 NEW YORK AVE., NW, SUITE 600

; CITY: WASHINGTON

; STATE: DC

; COUNTRY: USA

; ZIP: 20005

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/956,086

; FILING DATE: 20-Sep-2001

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/069,821

; FILING DATE: <Unknown>

; APPLICATION NUMBER: US 60/063,074

; FILING DATE: 27-OCT-1997

; APPLICATION NUMBER: US 60/050,472

; FILING DATE: 23-JUN-1997

; APPLICATION NUMBER: US 60/044,449

; FILING DATE: 30-APR-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: KIM, JUDITH U.

; REGISTRATION NUMBER: 40,679

; REFERENCE/DOCKET NUMBER: 0977.2280003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)371-2600

; TELEFAX: (202)371-2540

; INFORMATION FOR SEQ ID NO: 18:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: No. US20020155498A1 Relevant

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 18:

US-09-956-086-18 ;

Query Match 47.2%; Score 17; DB 9; Length 7;

Best Local Similarity 100.0%; Pred. No. 1.5e+05;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3
Db 3 TNN 5

RESULT 9

US-09-956-087-18

; Sequence 18, Application US/09956087

Patent No. US20020161201A1
GENERAL INFORMATION:
APPLICANT: FILPULA, DAVID
WANG, MAOLIANG
SHORR, ROBERT
WHITLOW, MARC
LEE, LIHSYNG S.
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/956,087
FILING DATE: 20-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/069,821
FILING DATE: 1998-04-30
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-956-087-18
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 3 TNN 5
RESULT 10
US-09-995-973-20
Sequence 20, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59

Patent No. US20020161201A1
GENERAL INFORMATION:
APPLICANT: FILPULA, DAVID
WANG, MAOLIANG
SHORR, ROBERT
WHITLOW, MARC
LEE, LIHSYNG S.
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/956,087
FILING DATE: 20-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/069,821
FILING DATE: 1998-04-30
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-956-087-18
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 3 TNN 5
RESULT 10
US-09-995-973-20
Sequence 20, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59

Patent No. US20020161201A1
GENERAL INFORMATION:
APPLICANT: FILPULA, DAVID
WANG, MAOLIANG
SHORR, ROBERT
WHITLOW, MARC
LEE, LIHSYNG S.
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/956,087
FILING DATE: 20-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/069,821
FILING DATE: 1998-04-30
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-956-087-18
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 3 TNN 5
RESULT 10
US-09-995-973-20
Sequence 20, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59

Patent No. US20020161201A1
GENERAL INFORMATION:
APPLICANT: FILPULA, DAVID
WANG, MAOLIANG
SHORR, ROBERT
WHITLOW, MARC
LEE, LIHSYNG S.
TITLE OF INVENTION: SINGLE-CHAIN ANTIGEN-BINDING PROTEINS
CAPABLE OF GLYCOSYLATION, PRODUCTION AND USES THEREOF
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVE., NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/956,087
FILING DATE: 20-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/069,821
FILING DATE: 1998-04-30
APPLICATION NUMBER: US 60/063,074
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: US 60/050,472
FILING DATE: 23-JUN-1997
APPLICATION NUMBER: US 60/044,449
FILING DATE: 30-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: KIM, JUDITH U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 0977.2280003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-956-087-18
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 3 TNN 5
RESULT 10
US-09-995-973-20
Sequence 20, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: zinc finger
OTHER INFORMATION: binding domain
US-09-995-973-20
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 1 TNN 3
RESULT 11
US-09-995-973-24
Sequence 24, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: zinc finger
OTHER INFORMATION: binding domain
US-09-995-973-24
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 1 TNN 3
RESULT 12
US-09-995-973-27
Sequence 27, Application US/09995973
Publication No. US20030024006A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: ULLMAN, Christopher G.
TITLE OF INVENTION: GENE SWITCHES
FILE REFERENCE: 8325-2003 / G7-US1
CURRENT APPLICATION NUMBER: US/09/995,973
CURRENT FILING DATE: 2002-03-19
NUMBER OF SEQ ID NOS: 59
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 27
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: zinc finger
OTHER INFORMATION: binding domain
US-09-995-973-27
Query Match 47.2%; Score 17; DB 9; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TNN 3
Db 1 TNN 3

Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3
|||
Db 1 TNN 3

RESULT 13

US-09-865-483-10
; Sequence 10, Application US/09865483
; Patent No. US20020061581A1
; GENERAL INFORMATION:
; APPLICANT: LEE, Jong Wook et al.
; TITLE OF INVENTION: A VARIABLE REGION OF THE MONOCLONAL ANTIBODY AGAINST THE HBV S-S
; TITLE OF INVENTION: ANTIGEN AND A GENE ENCODING THE SAME
; FILE REFERENCE: 1599-0197P
; CURRENT APPLICATION NUMBER: US/09/865,483
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-865-483-10

Query Match 47.2%; Score 17; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3
|||
Db 2 TNN 4

RESULT 14

US-09-924-703-12
; Sequence 12, Application-US/09924703
; Patent No. US20020137898A1
; GENERAL INFORMATION:
; APPLICANT: TRINCHIERI, GIORGIO
; APPLICANT: PERUSSIA, BICE
; APPLICANT: CLARK, STEVEN C.
; APPLICANT: WONG, GORDON G.
; APPLICANT: HEWICK, RODNEY
; APPLICANT: KOBAYASHI, MICHIO
; TITLE OF INVENTION: ANTIBODIES TO NATURAL KILLER STIMULATORY FACTOR
; FILE REFERENCE: 01142.0142.0100
; CURRENT APPLICATION NUMBER: US/09/924,703
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/325,958
; PRIOR FILING DATE: 1999-06-04
; PRIOR APPLICATION NUMBER: 08/858,000
; PRIOR FILING DATE: 1997-05-16
; PRIOR APPLICATION NUMBER: 08/403,013
; PRIOR FILING DATE: 1995-03-13
; PRIOR APPLICATION NUMBER: 07/584,941
; PRIOR FILING DATE: 1990-09-18
; PRIOR APPLICATION NUMBER: 07/307,817
; PRIOR FILING DATE: 1989-02-07
; PRIOR APPLICATION NUMBER: 07/269,945
; PRIOR FILING DATE: 1988-11-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-924-703-12

Query Match 47.2%; Score 17; DB 10; Length 8;
Best Local Similarity 60.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVLQ 6
:|||
Db 3 SNMLQ 7

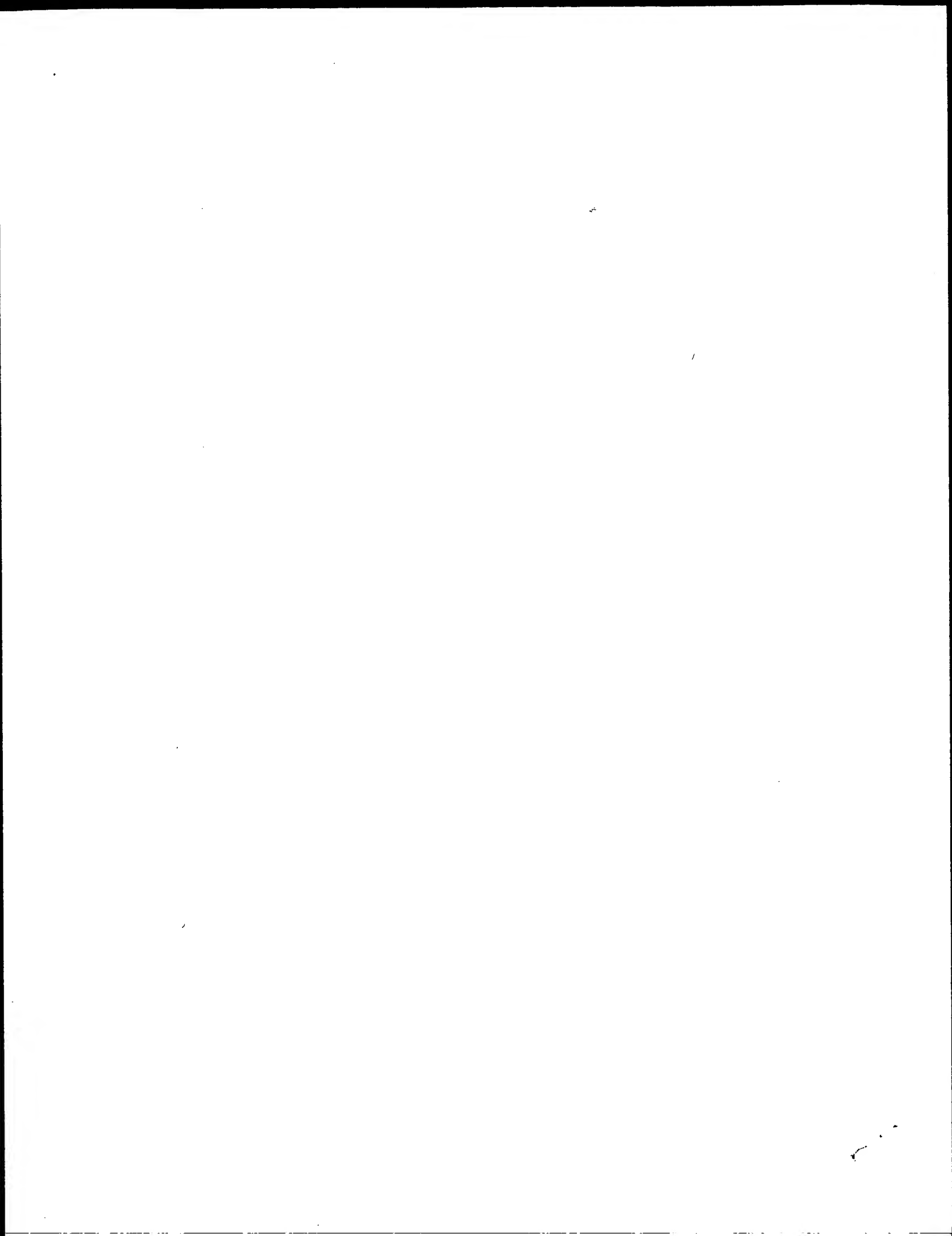
RESULT 15

US-09-990-762-54
; Sequence 54, Application US/09990762
; Patent No. US20020119498A1
; GENERAL INFORMATION:
; APPLICANT: JOUNG, J. KEITH
; APPLICANT: MILLER, JEFFREY
; APPLICANT: PABO, CARL O.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INTERACTION TRAP ASSAYS
; FILE REFERENCE: MTV-030.02 (20021-3002)
; CURRENT APPLICATION NUMBER: US/09/990,762
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 09/858,852
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/204,509
; PRIOR FILING DATE: 2000-05-16
; NUMBER OF SEQ ID NOS: 97
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Zinc finger
; OTHER INFORMATION: recognition sequence
US-09-990-762-54

Query Match 44.4%; Score 16; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNV 4
|||
Db 4 NNV 6

Search completed: February 26, 2003, 15:52:46
Job time : 13 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:43:32 ; Search time 346 Seconds
(without alignments) 14.907 Million cell updates/sec

Title: US-09-476-485A-24
Perfect score: 36
Sequence: 1 TNNVLQXT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 223113

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match Length	ID	Description	
1	34	94.4	8	1	PCT-US98-13046-24
2	34	94.4	8	18	US-09-476-485A-24
3	34	94.4	8	24	US-10-045-353-24
4	25	69.4	8	1	PCT-US01-08656-8073
5	22	61.1	6	22	US-09-823-823-61
6	22	61.1	6	22	US-09-823-823-61

7	20	55.6	7	10	US-08-697-360-22	Sequence 22, Appl
8	20	55.6	7	10	US-08-697-419-22	Sequence 22, Appl
9	20	55.6	7	10	US-08-697-440-22	Sequence 22, Appl
10	20	55.6	7	10	US-08-697-471-22	Sequence 22, Appl
11	20	55.6	7	10	US-08-697-471B-22	Sequence 22, Appl
12	20	55.6	7	10	US-08-697-472-22	Sequence 22, Appl
13	20	55.6	7	10	US-08-697-473-22	Sequence 22, Appl
14	20	55.6	7	11	US-08-702-105-22	Sequence 22, Appl
15	20	55.6	7	11	US-08-702-110-22	Sequence 22, Appl
16	20	55.6	7	11	US-08-702-168-22	Sequence 22, Appl
17	20	55.6	7	12	US-08-874-503-22	Sequence 22, Appl
18	20	55.6	7	22	US-09-823-823-80	Sequence 80, Appl
19	20	55.6	7	22	US-09-823-829-80	Sequence 80, Appl
20	20	55.6	7	22	US-09-848-885-22	Sequence 22, Appl
21	20	55.6	8	16	US-09-206-786-59	Sequence 59, Appl
22	20	55.6	8	16	US-09-206-786A-59	Sequence 59, Appl
23	20	55.6	8	18	US-09-458-297-113	Sequence 113, App
24	20	55.6	8	18	US-09-458-297-306	Sequence 306, App
25	20	55.6	8	18	US-09-458-297A-113	Sequence 113, App
26	20	55.6	8	18	US-09-458-297A-306	Sequence 306, App
27	19	52.8	7	1	PCT-US00-06588-36	Sequence 36, Appl
28	19	52.8	7	1	PCT-US00-06588-37	Sequence 37, Appl
29	19	52.8	7	20	US-09-637-216-1	Sequence 1, Appl
30	19	52.8	7	20	US-09-637-216-42	Sequence 42, Appl
31	19	52.8	7	20	US-09-637-216-44	Sequence 44, Appl
32	19	52.8	8	18	US-09-412-863-1681	Sequence 1681, Ap
33	19	52.8	8	18	US-09-412-863-4370	Sequence 4370, Ap
34	19	52.8	8	18	US-09-412-863-6754	Sequence 6754, Ap
35	18	50.0	5	8	US-08-406-779-89	Sequence 89, Appl
36	18	50.0	5	19	US-09-515-965A-1690	Sequence 1690, Ap
37	18	50.0	5	23	US-09-350-369C-547	Sequence 547, Appl
38	18	50.0	5	24	US-10-005-305-98	Sequence 98, Appl
39	18	50.0	6	13	US-08-984-098-106	Sequence 106, App
40	18	50.0	6	19	US-09-515-965A-1691	Sequence 1691, Ap
41	18	50.0	6	23	US-09-350-369C-548	Sequence 548, App
42	18	50.0	6	24	US-10-005-305-99	Sequence 99, Appl
43	18	50.0	7	8	US-08-407-620A-5	Sequence 5, Appl
44	18	50.0	7	19	US-09-515-965A-1692	Sequence 1692, Ap
45	18	50.0	7	20	US-09-637-216-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
PCT-US98-13046-24
; Sequence 24, Application PC/TUS9813046
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hofmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: USA
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/13046
; FILING DATE: June 23, 1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/881,189
; FILING DATE: June 24, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601

REFERENCE/DOCKET NUMBER: 381-44 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US98-13046-24

Query Match 94.4%; Score 34; DB 1; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
|||||
Db 1 TNNVLQVT 8

RESULT 2

US-09-476-485A-24
Sequence 24, Application US/09476485A
GENERAL INFORMATION:
APPLICANT: Colucci, M. Gabriella
APPLICANT: Chrispeels, Maarten J.
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
TITLE OF INVENTION: and Products of Their Use
FILE REFERENCE: 108236.119
CURRENT APPLICATION NUMBER: US/09/476,485A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 08/881,189
PRIOR FILING DATE: 1997-06-24
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin version 3.0
SEQ ID NO 24
LENGTH: 8
TYPE: PRT
ORGANISM: Dolichos lablab
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (7)..(7)
OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = any amino acid.
US-09-476-485A-24

Query Match 94.4%; Score 34; DB 18; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
|||||
Db 1 TNNVLQXT 8

RESULT 3

US-10-045-353-24
Sequence 24, Application US/10045353
GENERAL INFORMATION:
APPLICANT: Colucci et al.
TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
PROGENITOR CELL PRESERVATION FACTOR
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: USA
ZIP: 11753
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM compatible

OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/045,353
FILING DATE: 29-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/881,189
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Feit, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: 381-44 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-10-045-353-24

Query Match 94.4%; Score 34; DB 24; Length 8;
Best Local Similarity 87.5%; Pred. No. 4.2e+06;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLQXT 8
|||||
Db 1 TNNVLQVT 8

RESULT 4

PCT-US01-08656-8073
Sequence 8073, Application PC/TUS0108656
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066
CURRENT APPLICATION NUMBER: PCT/US01/08656
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO 8073
LENGTH: 8
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-08656-8073

Query Match 69.4%; Score 25; DB 1; Length 8;
Best Local Similarity 71.4%; Pred. No. 4.2e+06;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 NNVLQXT 8
|||||
Db 1 NNVLKFT 7

RESULT 5

US-09-823-823-61
Sequence 61, Application US/09823823
GENERAL INFORMATION:
APPLICANT: Yamamoto, Satoshi
APPLICANT: Kasai, Hiroaki
APPLICANT: Nakamura, Shoko
APPLICANT: Suzuki, Makoto
APPLICANT: Hamada, Tohru
TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS USING

; TITLE OF INVENTION: GENE AS AN INDICATOR
 ; FILE REFERENCE: 12817-004001
 ; CURRENT APPLICATION NUMBER: US/09/823.823
 ; CURRENT FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: US 09/208,688
 ; PRIOR FILING DATE: 1998-12-10
 ; PRIOR APPLICATION NUMBER: JP 97/343316
 ; PRIOR FILING DATE: 1997-12-12
 ; NUMBER OF SEQ ID NOS: 80
 ; SOFTWARE: PatentIn version 2.0
 ; SEQ ID NO 61
 ; LENGTH: 6
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetically generated protein
 US-09-823-823-61

Query Match 61.1%; Score 22; DB 22; Length 6;
 Best Local Similarity 66.7%; Pred. No. 4.2e+06;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 Db 1 TNNIPQ 6

RESULT 6

US-09-823-829-61
 ; Sequence 61, Application US/09823829
 ; GENERAL INFORMATION:
 ; APPLICANT: Yamamoto, Satoshi
 ; APPLICANT: Nakamura, Shoko
 ; APPLICANT: Suzuki, Makoto
 ; APPLICANT: Kasai, Hiroaki
 ; APPLICANT: Hamada, Tohru

; TITLE OF INVENTION: METHOD FOR IDENTIFICATION AND DETECTION OF MICROORGANISMS
 ; FILE REFERENCE: 12817-004001

; CURRENT APPLICATION NUMBER: US/09/823.829
 ; CURRENT FILING DATE: 2001-03-30
 ; PRIOR APPLICATION NUMBER: US 09/208,688
 ; PRIOR FILING DATE: 1998-12-10
 ; PRIOR APPLICATION NUMBER: JP 97/343316
 ; PRIOR FILING DATE: 1997-12-12
 ; NUMBER OF SEQ ID NOS: 82
 ; SOFTWARE: PatentIn version 2.0
 ; SEQ ID NO 61
 ; LENGTH: 6
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: synthetically generated peptide

US-09-823-829-61

Query Match 61.1%; Score 22; DB 22; Length 6;
 Best Local Similarity 66.7%; Pred. No. 4.2e+06;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
 Db 1 TNNIPQ 6

RESULT 7

US-08-697-360-22
 ; Sequence 22, Application US/08697360
 ; GENERAL INFORMATION:
 ; APPLICANT: Levitt, Roy C.
 ; APPLICANT: Maloy, W. Lee
 ; APPLICANT: Kari, Prasad
 ; APPLICANT: Nicolaides, Nicholas C.

; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating

; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/697,360
 ; FILING DATE: 23-AUG-1996
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fordis, Jean B.
 ; REGISTRATION NUMBER: 32,984
 ; REFERENCE/DOCKET NUMBER: 05387.0056-02000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202)408-4000
 ; TELEFAX: (202)408-4400
 ; INFORMATION FOR SEQ ID NO: 22:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-697-360-22

Query Match 55.6%; Score 20; DB 10; Length 7;
 Best Local Similarity 75.0%; Pred. No. 4.2e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
 Db 4 TNNI 7

RESULT 8

US-08-697-419-22
 ; Sequence 22, Application US/08697419
 ; GENERAL INFORMATION:
 ; APPLICANT: Levitt, Roy C.
 ; APPLICANT: Maloy, W. Lee
 ; APPLICANT: Kari, Prasad
 ; APPLICANT: Nicolaides, Nicholas C.

; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/697,419
 ; FILING DATE: 23-AUG-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fordis, Jean B.
 ; REGISTRATION NUMBER: 32,984

Query Match 55.6%; Score 20; DB 10; Length 7;
 Best Local Similarity 75.0%; Pred. No. 4.2e+06;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
 Db 4 TNNI 7

RESULT 8

US-08-697-419-22
 ; Sequence 22, Application US/08697419
 ; GENERAL INFORMATION:
 ; APPLICANT: Levitt, Roy C.
 ; APPLICANT: Maloy, W. Lee
 ; APPLICANT: Kari, Prasad
 ; APPLICANT: Nicolaides, Nicholas C.

; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
 ; STREET: 1300 I Street, N.W., Suite 700
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: IBM PC compatible
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/697,419
 ; FILING DATE: 23-AUG-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fordis, Jean B.
 ; REGISTRATION NUMBER: 32,984

REFERENCE/DOCKET NUMBER: 05387.0056-01000
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-697-419-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 9
US-08-697-440-22
; Sequence 22, Application US/08697440
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/697,440
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0056-09000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-697-440-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 10
US-08-697-471-22
; Sequence 22, Application US/08697471
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/697,471
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0056-05000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)408-4000
; TELEFAX: (202)408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-697-471-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 11
US-08-697-471B-22
; Sequence 22, Application US/08697471B
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
; TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
US-08-697-471B-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,471B
FILING DATE: 23-AUG-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0056-05000
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-697-471B-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 4 TNNI 7

RESULT 12
US-08-697-472-22
Sequence 22, Application US/08697472
GENERAL INFORMATION:
APPLICANT: Levitt, Roy C.
APPLICANT: Maloy, W. Lee
APPLICANT: Kari, Prasad
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,472
FILING DATE: 23-AUG-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0056-04000
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-697-472-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 4 TNNI 7

RESULT 13
US-08-697-473-22
Sequence 22, Application US/08697473
GENERAL INFORMATION:
APPLICANT: Levitt, Roy C.
APPLICANT: Maloy, W. Lee
APPLICANT: Kari, Prasad
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,473
FILING DATE: 23-AUG-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0056-03000
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-697-473-22

Query Match 55.6%; Score 20; DB 10; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
Db 4 TNNI 7

RESULT 14
US-08-702-105-22
Sequence 22, Application US/08702105
GENERAL INFORMATION:
APPLICANT: Levitt, Roy C.
APPLICANT: Maloy, W. Lee
APPLICANT: Kari, Prasad
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.

NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,105
FILING DATE: 23-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0056-06000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-105-22

Query Match 55.6%; Score 20; DB 11; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

RESULT 15
US-08-702-110-22
Sequence 22, Application US/08702110
GENERAL INFORMATION:
APPLICANT: Levitt, Roy C.
APPLICANT: Maloy, W. Lee
APPLICANT: Kari, Prasad
APPLICANT: Nicolaides, Nicholas C.
TITLE OF INVENTION: Asthma Associated Factors As Targets For Treating
TITLE OF INVENTION: Atopic Allergies Including Asthma And Related Disorders.
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,110
FILING DATE: 23-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B.
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0056-07000

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-110-22

Query Match 55.6%; Score 20; DB 11; Length 7;
Best Local Similarity 75.0%; Pred. No. 4.2e+06;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNNV 4
Db 4 TNNI 7

Search completed: February 26, 2003, 15:51:57
Job time : 347 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:43:47 ; Search time 23 seconds
(without alignments) 31.790 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 40840

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	94.4	8	6	US-10-190-258A-11
2	23	63.9	8	5	US-09-641-528A-4816
3	23	63.9	8	5	US-09-641-528A-8925
4	23	63.9	8	5	US-09-641-528A-13537
5	23	63.9	8	5	US-09-641-528A-16461
6	23	63.9	8	5	US-09-641-528A-18438
7	23	63.9	8	5	US-09-641-528A-19958
8	20	55.6	8	5	US-09-641-528A-21718
9	20	55.6	8	5	US-09-641-528A-30040
10	19	52.8	7	5	US-09-637-2168-1
11	19	52.8	7	5	US-09-637-2168-42
12	19	52.8	7	5	US-09-637-2168-44
13	18	50.0	7	1	PCT-US02-35009-11
14	18	50.0	7	5	US-09-637-2168-33
15	18	50.0	7	5	US-10-284-252-11
16	18	50.0	7	5	US-09-641-528A-4696
17	18	50.0	8	5	US-09-641-528A-8925
18	18	50.0	8	5	US-09-641-528A-13537
19	18	50.0	8	5	US-09-641-528A-16461
20	18	50.0	8	5	US-09-641-528A-18438
21	18	50.0	8	5	US-09-641-528A-19958
22	18	50.0	8	5	US-09-641-528A-21718
23	18	50.0	8	5	US-09-641-528A-30040
24	18	50.0	8	5	US-09-641-528A-41860
25	18	50.0	8	5	US-09-641-528A-4696
26	18	50.0	8	5	US-09-641-528A-4816

```

27 18 50.0 8 5 US-09-641-528A-4816 Sequence 4816, Ap
28 18 50.0 8 5 US-09-641-528A-8925 Sequence 8925, Ap
29 18 50.0 8 5 US-09-641-528A-13537 Sequence 13537, A
30 18 50.0 8 5 US-09-641-528A-16461 Sequence 16461, A
31 18 50.0 8 5 US-09-641-528A-18438 Sequence 18438, A
32 18 50.0 8 5 US-09-641-528A-19958 Sequence 19958, A
33 18 50.0 8 5 US-09-641-528A-21718 Sequence 21718, A
34 18 50.0 8 5 US-09-641-528A-30040 Sequence 30040, A
35 18 50.0 8 5 US-09-641-528A-41860 Sequence 41860, A
36 17 47.2 5 5 US-09-939-481-36 Sequence 36, Appl
37 17 47.2 5 5 US-09-939-481-37 Sequence 37, Appl
38 17 47.2 5 5 US-09-939-481-38 Sequence 38, Appl
39 17 47.2 5 6 US-10-237-850-36 Sequence 36, Appl
40 17 47.2 5 6 US-10-237-850-37 Sequence 37, Appl
41 17 47.2 5 6 US-10-237-850-38 Sequence 38, Appl
42 17 47.2 5 6 US-09-637-2168-37 Sequence 37, Appl
43 17 47.2 7 6 US-10-258-144-160 Sequence 160, App
44 17 47.2 7 6 US-10-053-498B-267 Sequence 267, App
45 17 47.2 8 5 US-09-641-528-5920 Sequence 5920, Ap

```

ALIGNMENTS

```

RESULT 1
US-10-190-258A-11
; Sequence 11, Application US/10190258A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G
; TITLE OF INVENTION: Dendritic Cell Isolation Methods
; FILE REFERENCE: 108236.132
; CURRENT APPLICATION NUMBER: US/10/190,258A
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/303,265
; PRIOR FILING DATE: 2001-07-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Dolichos lablab
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 7
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-190-258A-11

```

```

Query Match          94.4%; Score 34; DB 6; Length 8;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TNNVLOXT 8
   |||||
Db 1 TNNVLOXT 8

```

```

RESULT 2
US-09-641-528-4836
; Sequence 4836, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

```

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4836
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-4836

Query Match 63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| :||
Db 2 TRNLIQ 7

RESULT 3

US-09-641-528-18778
; Sequence 18778, Application US/09641528
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 18778
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-18778

Query Match 63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| :||
Db 1 TRNLIQ 6

RESULT 4

US-09-641-528-21808

; Sequence 21808, Application US/09641528
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 21808
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-21808

Query Match 63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| :||
Db 2 TRNLIQ 7

RESULT 5

US-09-641-528A-4836

; Sequence 4836, Application US/09641528A
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 4836
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-4836

Query Match 63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TNNVLQ 6
| :||
Db 2 TRNLIQ 7

RESULT 6

US-09-641-528A-18778

; Sequence 18778, Application US/09641528A
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0


```
; SEQ ID NO 18778
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-18778

Query Match          63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 TNNVLQ 6
   | | : | |
Db 1 TRNIIQ 6

RESULT 7
US-09-641-528A-21808
; Sequence 21808, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641.528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21808
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-21808

Query Match          63.9%; Score 23; DB 5; Length 8;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 TNNVLQ 6
   | | : | |
Db 2 TRNIIQ 7

RESULT 8
US-09-641-528-44907
; Sequence 44907, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641.528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44907
```

```
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-44907

Query Match          55.6%; Score 20; DB 5; Length 8;
Best Local Similarity 60.0%; Pred. No. 4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVILQ 6
   | | | : |
Db 1 NNVIE 5

RESULT 9
US-09-641-528A-44907
; Sequence 44907, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641.528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44907
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-44907

Query Match          55.6%; Score 20; DB 5; Length 8;
Best Local Similarity 60.0%; Pred. No. 4e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 NNVILQ 6
   | | | : |
Db 1 NNVIE 5

RESULT 10
US-09-637-216B-1
; Sequence 1, Application US/09637216B
; GENERAL INFORMATION:
; APPLICANT: WASHINGTON UNIVERSITY
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF
; FILE REFERENCE: WSHU2005.1
; CURRENT APPLICATION NUMBER: US/09/637,216B
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,280
; PRIOR FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized
```

; OTHER INFORMATION: Sequence
US-09-637-216B-1

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

RESULT 11

US-09-637-216B-42
; Sequence 42, Application US/09637216B
; GENERAL INFORMATION:
; APPLICANT: WASHINGTON UNIVERSITY
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF
; FILE REFERENCE: WSHU2005.1
; CURRENT APPLICATION NUMBER: US/09/637,216B
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,280
; PRIOR FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized
; OTHER INFORMATION: Sequence
US-09-637-216B-42

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

RESULT 12

US-09-637-216B-44
; Sequence 44, Application US/09637216B
; GENERAL INFORMATION:
; APPLICANT: WASHINGTON UNIVERSITY
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF
; FILE REFERENCE: WSHU2005.1
; CURRENT APPLICATION NUMBER: US/09/637,216B
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,280
; PRIOR FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized
; OTHER INFORMATION: Sequence
US-09-637-216B-44

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

Query Match 52.8%; Score 19; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

RESULT 13

PCT-US02-35009-11
; Sequence 11, Application PC/TUS0235009
; GENERAL INFORMATION:
; APPLICANT: Jones, C. Hal
; APPLICANT: Dexter, Paul L.
; APPLICANT: Evans, Amy K.
; APPLICANT: Hruby, Dennis E.
; TITLE OF INVENTION: DEGP Protease: Cleavage Site
; TITLE OF INVENTION: Identification and Proteolysis of a Natural Target in E.
; TITLE OF INVENTION: Coli
; FILE REFERENCE: 016921-169
; CURRENT APPLICATION NUMBER: PCT/US02/35009
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: US 60/330,855
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
PCT-US02-35009-11

Query Match 50.0%; Score 18; DB 1; Length 7;
Best Local Similarity 56.7%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NVLQ 8
| | | |
Db 1 NVLQ 6

RESULT 14

US-09-637-216B-33
; Sequence 33, Application US/09637216B
; GENERAL INFORMATION:
; APPLICANT: WASHINGTON UNIVERSITY
; TITLE OF INVENTION: ANTI-BACTERIAL COMPOUNDS DIRECTED AGAINST PILUS
; TITLE OF INVENTION: BIOGENESIS, ADHESION AND ACTIVITY; CO-CRYSTALS OF PILUS
; TITLE OF INVENTION: SUBUNITS AND METHODS OF USE THEREOF
; FILE REFERENCE: WSHU2005.1
; CURRENT APPLICATION NUMBER: US/09/637,216B
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,280
; PRIOR FILING DATE: 1999-08-11
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthesized
; OTHER INFORMATION: Sequence
US-09-637-216B-33

Query Match 50.0%; Score 18; DB 5; Length 7;
Best Local Similarity 75.0%; Pred. No. 4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

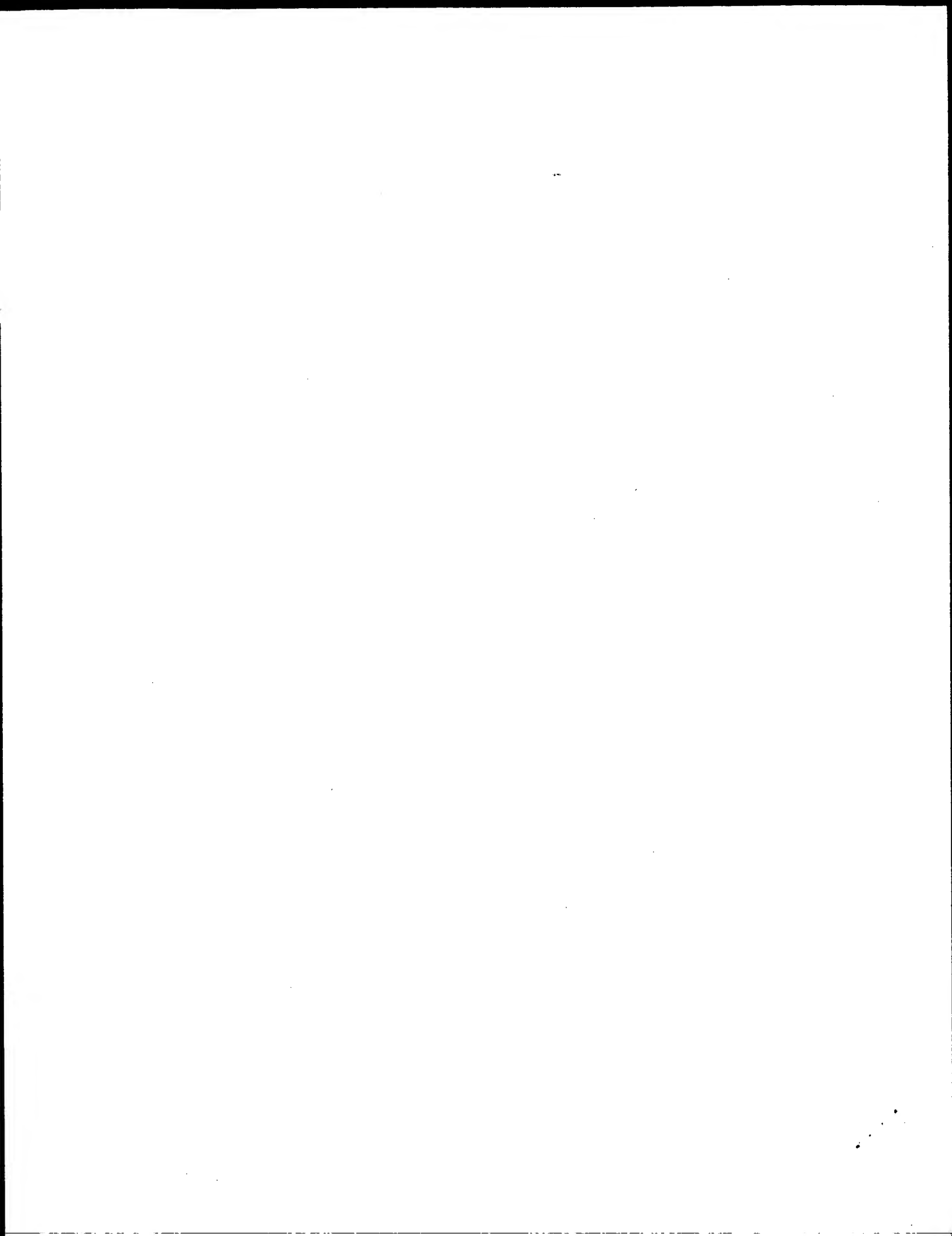
QY 3 NVLQ 6
| | | |
Db 1 NVLQ 4

RESULT 15
US-10-284-252-11
; Sequence 11, Application US/10284252
; GENERAL INFORMATION:
; APPLICANT: Jones, C. Hal
; APPLICANT: Dexter, Paul L.
; APPLICANT: Evans, Amy K.
; APPLICANT: Hruby, Dennis E.
; TITLE OF INVENTION: DEGP Protease: Cleavage Site
; TITLE OF INVENTION: Identification and Proteolysis of a Natural Target in E.
; TITLE OF INVENTION: Coll
; FILE REFERENCE: 016921-169
; CURRENT APPLICATION NUMBER: US/10/284,252
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/330,855
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-10-284-252-11

Query Match 50.0%; Score 18; DB 6; Length 7;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 NVLQXT 8
 111 1
Db 1 NVLHYT 6

Search completed: February 26, 2003, 15:52:27
Job time : 23 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 15:42:51 ; Search time 14 seconds
(without alignments)
16.813 Million cell updates/sec

Title: US-09-476-485A-24

Perfect score: 36

Sequence: 1 TNNVLOXT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 53074

Minimum DB seq length: 0

Maximum DB seq length: 8

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA.*
- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pap.*
 - 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pap.*
 - 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pap.*
 - 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pap.*
 - 5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pap.*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	94.4	8	4	US-08-881-189B-24
2	20	55.6	7	2	US-08-702-105A-22
3	20	55.6	7	3	US-08-702-110A-22
4	20	55.6	7	4	US-09-325-571-22
5	18	50.0	6	2	US-07-662-764D-25
6	17	47.2	5	3	US-08-915-189-36
7	17	47.2	5	3	US-08-915-189-37
8	17	47.2	5	3	US-08-915-189-38
9	17	47.2	5	3	US-08-972-760-36
10	17	47.2	5	3	US-08-972-760-37
11	17	47.2	5	3	US-08-972-760-38
12	17	47.2	5	4	US-09-089-645A-36
13	17	47.2	5	4	US-09-089-645A-37
14	17	47.2	5	4	US-09-089-645A-38
15	17	47.2	5	4	US-09-503-998-36
16	17	47.2	5	4	US-09-503-998-37
17	17	47.2	5	4	US-09-503-998-38
18	17	47.2	7	1	US-08-346-333-32
19	17	47.2	7	2	US-08-392-973A-34
20	17	47.2	7	2	US-08-946-241B-11
21	17	47.2	7	3	US-09-309-053-11
22	17	47.2	7	4	US-09-069-821-18
23	17	47.2	7	5	PCT-US91-07506-32
24	17	47.2	7	6	5210075-58
25	17	47.2	8	1	US-08-165-038-25
26	17	47.2	8	2	US-08-876-781-25
27	17	47.2	8	2	US-09-016-366A-65

28 16 44.4 4 1 US-08-456-424-48
29 16 44.4 5 4 US-08-974-549A-81
30 16 44.4 5 4 US-08-974-549A-97
31 16 44.4 6 1 US-08-486-721A-19
32 16 44.4 7 1 US-07-968-781A-52
33 16 44.4 7 1 US-08-486-721A-5
34 16 44.4 7 1 US-08-486-721A-10
35 16 44.4 7 1 US-08-486-721A-12
36 16 44.4 7 2 US-08-874-678-11
37 16 44.4 7 3 US-08-643-839-11
38 16 44.4 7 4 US-09-410-025-11
39 16 44.4 7 4 US-09-348-886-11
40 16 44.4 8 1 US-08-279-906A-12
41 16 44.4 8 1 US-08-279-906A-13
42 16 44.4 8 1 US-08-486-721A-6
43 16 44.4 8 4 US-08-953-033-22
44 16 44.4 8 4 US-09-139-802-157
45 15 41.7 4 4 US-09-177-249-231

ALIGNMENTS

RESULT 1
US-08-881-189B-24
; Sequence 24, Application US/08881189B
; Patent No. 6310195
; GENERAL INFORMATION:
; APPLICANT: Colucci et al.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A LECTIN-DERIVED
; PROGENITOR CELL PRESERVATION FACTOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/881,189B
; FILING DATE: June 24, 1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Felt, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: 381-44 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-881-189B-24

Query Match 94.4%; Score 34; DB 4; Length 8;
Best Local Similarity 87.5%; Pred. No. 2e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TNNVLOXT 8

Db 1 TNNVLOXT 8

```

RESULT 2
US-08-702-105A-22
; Sequence 22, Application US/08702105A
; Patent No. 5908839
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For
; TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,105A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-702-105A-22

Query Match 55.6%; Score 20; DB 2; Length 7;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
   |||
Db 4 TNNI 7

RESULT 3
US-08-702-110A-22
; Sequence 22, Application US/08702110A
; Patent No. 6037149
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For
; TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,110A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-702-110A-22

Query Match 55.6%; Score 20; DB 3; Length 7;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNNV 4
   |||
Db 4 TNNI 7

RESULT 4
US-09-325-571-22
; Sequence 22, Application US/09325571
; Patent No. 6261559
; GENERAL INFORMATION:
; APPLICANT: Levitt, Roy C.
; APPLICANT: Maloy, W. Lee
; APPLICANT: Kari, U. Prasad
; APPLICANT: Nicolaides, Nicholas C.
; TITLE OF INVENTION: Asthma Associated Factors As Targets For
; TITLE OF INVENTION: Treating Atopic Allergies Including Asthma And Related
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner L.L.P.
; STREET: 1300 I Street N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/325,571
; FILING DATE:

```

```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/874,503
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32984
; REFERENCE/DOCKET NUMBER: 05387.0056-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-325-571-22

Query Match 55.6%; Score 20; DB 4; Length 7;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

QY 1 TNNV 4
Db 4 TNNI 7

;
; RESULT 5
; US-07-662-764D-25
; Sequence 25, Application US/07662764D
; Patent No. 5866363
; GENERAL INFORMATION:
; APPLICANT: Piecznik, George
; TITLE OF INVENTION: METHOD AND MEANS FOR SORTING AND
; TITLE OF INVENTION: IDENTIFYING BIOLOGICAL INFORMATION
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LERNER, DAVID, LITTENBERG, KRUMHOLZ &
; STREET: 600 South, Avenue West
; CITY: Westfield
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07090
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,764D
; FILING DATE: 28-FEB-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/201,358
; FILING DATE: 26-MAY-1988
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; ADDRESSEE: Medlen & Carroll, LLP

;
; MOLECULE TYPE: protein
; US-07-662-764D-25

Query Match 50.0%; Score 18; DB 2; Length 6;
Best Local Similarity 75.0%; Pred. No. 2e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0;

QY 3 NVLQ 6
Db 1 NILQ 4

;
; RESULT 6
; US-08-915-189-36
; Sequence 36, Application US/08915189
; Patent No. 6001965
; Patent No. 6001965 6001955
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Anticancer Compounds and Methods
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,189
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02877
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-915-189-36

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TNN 3
Db 3 TNN 5

;
; RESULT 7
; US-08-915-189-37
; Sequence 37, Application US/08915189
; Patent No. 6001965
; Patent No. 6001965 6001955
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Anticancer Compounds and Methods
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
```

; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,189
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02877
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-915-189-37

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
Db 3 TNN 5

RESULT 8
US-08-915-189-38
; Sequence 38, Application US/08915189
; Patent No. 6001965
; Patent No. 6001965 6001955
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Anticancer Compounds and Methods
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915,189
; FILING DATE: 20-AUG-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-02877
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-915-189-38

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
Db 3 TNN 5

RESULT 9
US-08-972-760-36
; Sequence 36, Application US/08972760
; Patent No. 6025150
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Methods and Compositions for Wound
; TITLE OF INVENTION: Healing
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,760
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/754,322
; FILING DATE: 21-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03057
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-972-760-36

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
Db 3 TNN 5

RESULT 10
US-08-972-760-37
; Sequence 37, Application US/08972760
; Patent No. 6025150

GENERAL INFORMATION:
APPLICANT: Livant, Donna L
TITLE OF INVENTION: Methods and Compositions for Wound
HEALING
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,760
FILING DATE: 18-NOV-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/754,322
FILING DATE: 21-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03057
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-972-760-37

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TNN 3
Db 3 TNN 5

RESULT 11
US-08-972-760-38
Sequence 38, Application US/08972760
Patent No. 6025150
GENERAL INFORMATION:
APPLICANT: Livant, Donna L
TITLE OF INVENTION: Methods and Compositions for Wound
HEALING
NUMBER OF SEQUENCES: 85
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,760
FILING DATE: 18-NOV-1997

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/754,322
FILING DATE: 21-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03057
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-08-972-760-38

Query Match 47.2%; Score 17; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0;

QY 1 TNN 3
Db 3 TNN 5

RESULT 12
US-09-089-645A-36
Sequence 36, Application US/09089645A
Patent No. 6140068
GENERAL INFORMATION:
APPLICANT: Livant, Donna L
TITLE OF INVENTION: Protease Resistant Compositions for
Wound Healing
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Medlen & Carroll, LLP
STREET: 220 Montgomery Street, Suite 2200
CITY: San Francisco
STATE: California
COUNTRY: United States Of America
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/089,645A
FILING DATE: 03-JUN-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/754,322
FILING DATE: 21-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/972,760
FILING DATE: 18-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03349
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant

; MOLECULE TYPE: peptide
US-09-089-645A-36

Query Match 47.2%; Score 17; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
|||
Db 3 TNN 5

RESULT 13

US-09-089-645A-37
; Sequence 37, Application US/09089645A
; Patent No. 6140068
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Protease Resistant Compositions for
; Wound Healing
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,645A
; FILING DATE: 03-JUN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/754,322
; FILING DATE: 21-NOV-1996

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/972,760
; FILING DATE: 18-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-089-645A-37

Query Match 47.2%; Score 17; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
|||
Db 3 TNN 5

RESULT 14

US-09-089-645A-38
; Sequence 38, Application US/09089645A
; Patent No. 6140068
; GENERAL INFORMATION:

; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Protease Resistant Compositions for
; Wound Healing
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/089,645A
; FILING DATE: 03-JUN-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/754,322
; FILING DATE: 21-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/972,760
; FILING DATE: 18-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-09-089-645A-38

Query Match 47.2%; Score 17; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TNN 3
|||
Db 3 TNN 5

RESULT 15

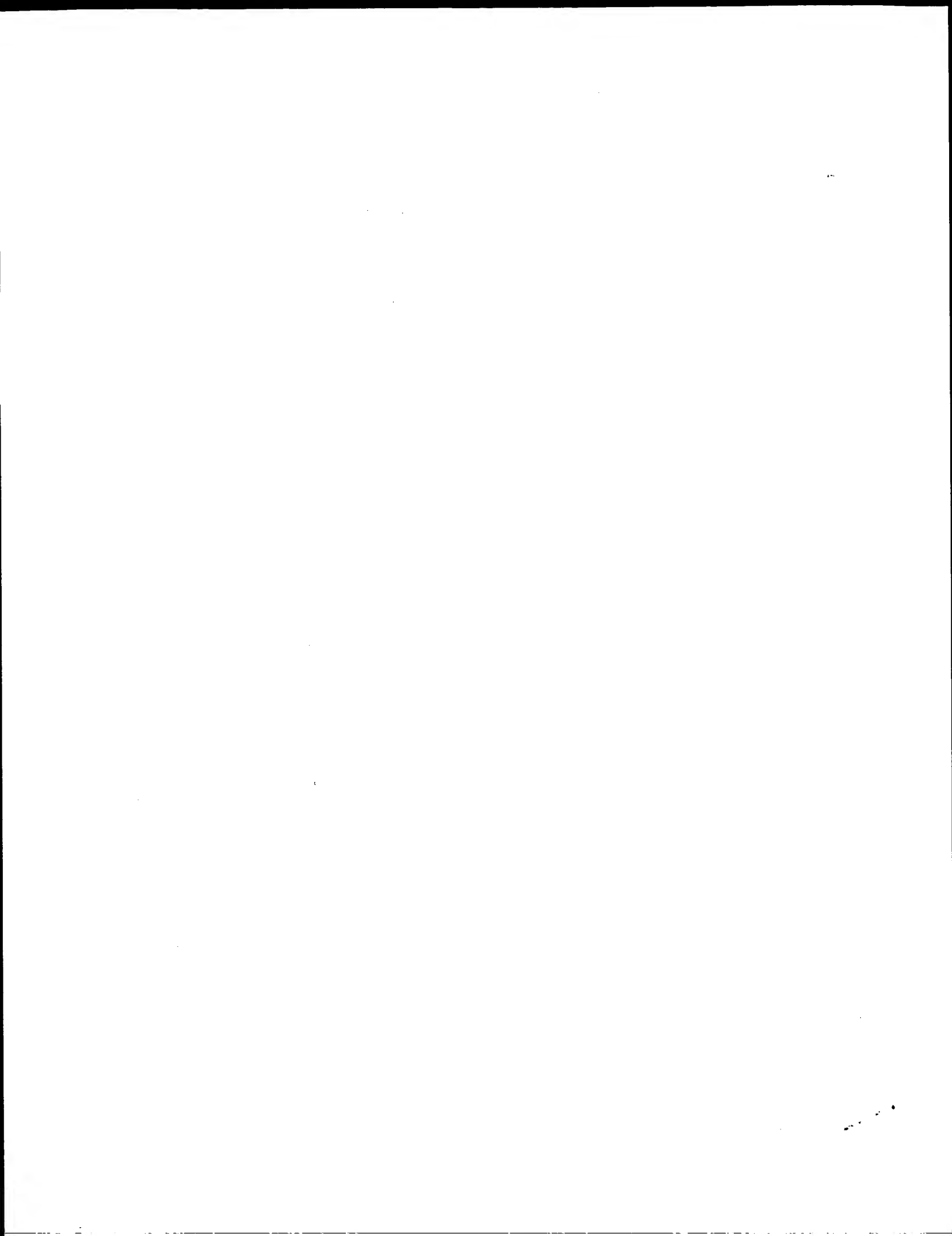
US-09-503-998-36
; Sequence 36, Application US/09503998
; Patent No. 6331409
; GENERAL INFORMATION:
; APPLICANT: Livant, Donna L
; TITLE OF INVENTION: Methods and Compositions for Wound
; Healing
; NUMBER OF SEQUENCES: 85
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States Of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/503,998
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/972,760
FILING DATE: 18-NOV-1997
APPLICATION NUMBER: US 08/754,322
FILING DATE: 21-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Carroll, Peter G.
REGISTRATION NUMBER: 32,837
REFERENCE/DOCKET NUMBER: UM-03057
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: No. 6331409 Relevant
TOPOLOGY: No. 6331409 Relevant
MOLECULE TYPE: peptide
US-09-503-998-36

Query Match 47.2%; Score 17; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TNN 3
Db 3 TNN 5

Search completed: February 26, 2003, 15:46:03
Job time : 15 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:55:30 ; Search time 45 Seconds
(without alignments)
29,908 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 2093

Minimum DB seq length: 0
Maximum DB seq length: 14

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	40.9	14	2 PA0007	lectin B1 - Psophocarpus
2	25	37.9	12	2 B61497	seed protein ws-17
3	22	33.3	10	2 C39572	sperm-activating p
4	22	33.3	14	2 C59137	protein Pf3 - gold
5	21	31.8	14	2 F61497	seed protein ws-21
6	21	31.8	14	2 A35105	hypothetical prote
7	21	31.8	14	2 B61597	cytochrome P450 AL
8	20	30.3	12	2 S26547	T-cell receptor be
9	19	28.8	14	2 B44854	L-2,4-diaminobuty
10	18	27.3	9	2 PD0443	3-oxoacid CoA-tran
11	18	27.3	10	2 S38305	lectin GML2 alpha
12	18	27.3	12	2 A61360	vespakinin M - hor
13	18	27.3	13	2 PQ0445	urotensin II - lau
14	18	27.3	13	2 PS0443	potassium channel
15	18	27.3	14	2 PT0077	proteochondroitin c
16	18	27.3	14	2 S07768	soluble hydrogenas
17	17.5	26.5	13	2 S23640	ig kappa chain J s
18	17	25.8	10	1 ECU04M	tachykinin IV - mi
19	17	25.8	10	2 C60788	sperm-activating p
20	17	25.8	10	2 A60787	sperm-activating p
21	17	25.8	10	2 A60527	sperm-activating p
22	17	25.8	10	2 I60527	sperm-activating p
23	17	25.8	10	2 A24867	scyllorhizin I - s
24	17	25.8	10	2 D61440	polygalacturonase
25	17	25.8	13	2 A33660	osteoclast functio
26	17	25.8	13	2 S10562	zona pellucida-bin
27	17	25.8	13	2 I77387	AMP deaminase - ra
28	17	25.8	14	2 S39932	S-allele-associate
29	17	25.8	14	2 B28018	very late antigen-

ALIGNMENTS

RESULT 1

PA0007
lectin B1 - Psophocarpus scandens (fragment)
C:Species: Psophocarpus scandens
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Mar-1995
C:Accession: PA0007
R:Kortt, A.A.

Phytochemistry 27, 2847-2855, 1988

A:Title: Isolation and characterization of the lectins from the seeds of Psophocarpus
A:Reference number: PA0005
A:Accession: PA0007
A:Molecule type: protein
A:Residues: 1-14 <KOR>
A:Experimental source: seed
A:Comment: The seeds of Psophocarpus contain two distinct groups of lectins which can
C:Keywords: lectin

Query Match 40.9%; Score 27; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 64;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Oy 2 QSLSFXTKFK 11
:::| | | |
Db 1 ETISFNFNQF 10

RESULT 2

B61497
seed protein ws-17 - winged bean (fragment)

C:Species: Psophocarpus tetragonolobus (winged bean)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
C:Accession: B61497
R:Hirano, H.

J. Protein Chem. 8, 115-130, 1989

A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-
A:Reference number: A61491; PMID:89351606; PMID:2765119
A:Accession: B61497

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-12 <HIR>

C:Keywords: seed

Query Match 37.9%; Score 25; DB 2; Length 12;
Best Local Similarity 44.4%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 3 SLSFXTKFK 11
:::| | | |
Db 2 TISFNFNQF 10

RESULT 3

C39572

sperm-activating peptide TG-3 - sea urchin (Tripneustes gratilla)
 N:Alternate names: Spermact homolog TG-3
 C:Species: Tripneustes gratilla
 C:Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 18-Aug-2000
 C:Accession: C39572
 R:Yoshino, K.; Takao, T.; Suhara, M.; Kitai, T.; Hori, H.; Nomura, K.; Yamaguchi, M.; Shibata, K.
 Biochemistry 30, 6203-6209, 1991
 A:Title: Identification of a novel amino acid, o-bromo-L-phenylalanine, in egg-associated proteins
 A:Reference number: A39572; MUID:91283461; PMID:2059627
 A:Accession: C39572
 A:Molecule type: protein
 A:Residues: 1-10 <YOS>
 C:Superfamily: unassigned animal peptides
 C:Keywords: bromine
 F:2/Modified site: 2'-bromophenylalanine (Phe) #status experimental

Query Match 33.3%; Score 22; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FDLD 14
 ||||
 Db 2 FDLD 5

RESULT 4

C59137
 Protein Pf3 - golden needle mushroom (fragment)
 C:Species: Flammulina velutipes (golden needle mushroom)
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: C59137
 R:Sakamoto, Y.; Ando, A.; Tamai, Y.; Miura, K.
 submitted to the Protein Sequence Database, November 1999
 A:Description: Differences of proteins expressed in the fruiting dikaryon and the non-fructifying dikaryon
 A:Reference number: A59137
 A:Accession: C59137
 A:Molecule type: protein
 A:Residues: 1-14 <SAK>

Query Match 33.3%; Score 22; DB 2; Length 14;
 Best Local Similarity 33.3%; Pred. No. 5.7e+02;
 Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 QSLSPFTKFDL 13
 ||||
 Db 2 QVLGPKYVFSI 13

RESULT 5

F61497
 seed protein ws-21 - winged bean (fragment)
 C:Species: Psophocarpus tetragonolobus (winged bean)
 C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
 C:Accession: F61497
 R:Hirano, H.
 J. Protein Chem. 8, 115-130, 1989
 A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-dimensional gel electrophoresis
 A:Reference number: A61491; MUID:89351606; PMID:2765119
 A:Accession: F61497
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-14 <HIR>
 C:Keywords: seed

Query Match 31.8%; Score 21; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 8.8e+02;
 Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SPXFTKFD 12
 ||||
 Db 5 SFNFXFE 12

RESULT 6

A35105
 hypothetical protein - Neurospora crassa mitochondrion (fragment)
 C:Species: mitochondrion Neurospora crassa
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Dec-1999
 C:Accession: A35105
 R:Saville, B.J.; Collins, R.A.
 Cell 61, 685-696, 1990
 A:Title: A site-specific self-cleavage reaction performed by a novel RNA in Neurospora crassa
 A:Reference number: A35105; MUID:90263093; PMID:2160856

A:Accession: A35105
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-14 <SAV>
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC3
 C:Keywords: mitochondrion

Query Match 31.8%; Score 21; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 8.8e+02;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 LSPXFTKFDL 13
 ||||
 Db 1 LSLTLTLQL 10

RESULT 7

B61597
 cytochrome P450 AL-2 - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
 C:Accession: B61597
 R:Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.
 Drug Metab. Dispos. 19, 291-297, 1991
 A:Title: Purification and aminopyrine monooxygenase activity of liver microsomal cytochrome P450 AL-2
 A:Reference number: A61597; MUID:91292910; PMID:1676625

A:Accession: B61597
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-14 <SHI>

Query Match 31.8%; Score 21; DB 2; Length 14;
 Best Local Similarity 55.6%; Pred. No. 8.8e+02;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 SLSPXFTK 11
 ||||
 Db 6 SLSPLLVGF 14

RESULT 8

S26547
 T-cell receptor beta chain (clone Cw3/10.1) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
 C:Accession: S26547
 R:Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; J. Exp. Med. 176, 439-447, 1992
 A:Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
 A:Reference number: S26512; MUID:92364546; PMID:1380061

A:Accession: S26547
 A:Molecule type: mRNA
 A:Residues: 1-12 <CAS>
 A:Cross-references: EMBL:X67997
 A:Experimental source: cytolytic T-lymphocyte, clone Cw3/10.1
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: T-cell receptor

Query Match 30.3%; Score 20; DB 2; Length 12;
 Best Local Similarity 44.4%; Pred. No. 1.2e+03;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 AQSLSFXFT 9
 | | | | |
 Db 2 ASSTGFDYT 10

RESULT 9

B44854
 L-2,4-diaminobutyrate decarboxylase (EC 4.1.1.-) - Vibrio alginolyticus (fragment)
 C:Species: Vibrio alginolyticus
 C:Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 10-Nov-1995
 C:Accession: B44854; B41817
 R:Yamamoto, S.; Tsuzaki, Y.; Tougon, K.; Shinoda, S.
 J. Gen. Microbiol. 138, 1461-1465, 1992
 A:Title: Purification and characterization of L-2,4-diaminobutyrate decarboxylase from
 A:Reference number: A44854; MUID:92381494; PMID:1512577
 A:Accession: B44854
 A:Molecule type: protein
 A:Residues: 1-14 <YAM>
 A>Note: sequence extracted from NCBI backbone (NCBIP:112332)
 C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 28.8%; Score 19; DB 2; Length 14;
 Best Local Similarity 50.0%; Pred. No. 2.1e+03;
 Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 TKFDLD 14
 | | | | |
 Db 2 TAFEVD 7

RESULT 10

PD0443
 3-oxoacid CoA-transferase (EC 2.8.3.5) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 05-Feb-1999
 C:Accession: PD0443
 R:Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.
 submitted to JIPID, August 1998
 A:Description: Proteome analysis of mouse brain.
 A:Reference number: PD0441
 A:Contents: Striatum
 A:Accession: PD0443
 A:Molecule type: protein
 A:Residues: 1-9 <KAM>
 C:Keywords: CoA-transferase

Query Match 27.3%; Score 18; DB 2; Length 9;
 Best Local Similarity 66.7%; Pred. No. 2.8e+05;
 Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 TKFDLD 14
 | | | | |
 Db 1 TKFYTD 6

RESULT 11

S38305
 lectin GNL2 alpha chain - kidney bean (fragment)
 C:Species: Phaseolus vulgaris (kidney bean)
 C:Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
 C:Accession: S38305
 R:Kamenura, K.; Furuichi, Y.; Umekawa, H.; Takahashi, T.
 Biochim. Biophys. Acta 1158, 181-188, 1993
 A:Title: Purification and characterization of novel lectins from Great Northern bean, Ph
 A:Reference number: S38304; MUID:94002183; PMID:8399319
 A:Accession: S38305
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <KAM>

Query Match 27.3%; Score 18; DB 2; Length 10;

Best Local Similarity 50.0%; Pred. No. 2.3e+03;
 Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFX 8
 | | | | |
 Db 1 ATETSFSS 8

RESULT 12

A61360
 vespakinin M - hornet (Vespa mandarinia)
 C:Species: Vespa mandarinia
 C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
 C:Accession: A61360
 R:Kishimura, H.; Yasuhara, T.; Yoshida, H.; Nakajima, T.
 Chem. Pharm. Bull. 24, 2896-2897, 1976
 A:Title: Vespakinin-M, a novel bradykinin analogue containing hydroxyproline, in the
 A:Reference number: A61360; MUID:77114342; PMID:1017116
 A:Accession: A61360
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-12 <KIS>
 C:Superfamily: unassigned animal peptides
 C:Keywords: hydroxyproline; venom
 F:4/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 27.3%; Score 18; DB 2; Length 12;
 Best Local Similarity 42.9%; Pred. No. 2.8e+03;
 Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 8 FTRFDLD 14
 | | | | |
 Db 6 FSPFRID 12

RESULT 13

PQ0445
 urotensin II - laughing frog
 C:Species: Rana ridibunda (laughing frog)
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Mar-1995
 C:Accession: PQ0445
 R:Conlon, J.M.; O'Harte, F.; Smith, D.D.; Tonon, M.C.; Vaudry, H.
 Biochem. Biophys. Res. Commun. 188, 578-583, 1992
 A:Title: Isolation and primary structure of urotensin II from the brain of a tetrapod
 A:Reference number: PQ0445; MUID:93075134; PMID:1445302
 A:Accession: PQ0445
 A:Molecule type: protein
 A:Residues: 1-13 <CON>
 A:Experimental source: brain
 C:Superfamily: urotensin II

Query Match 27.3%; Score 18; DB 2; Length 13;
 Best Local Similarity 45.5%; Pred. No. 3e+03;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKF 11
 | | | | |
 Db 1 AGNLSECFWKY 11

RESULT 14

PS0443
 potassium channel protein slo G3 - fruit fly (Drosophila melanogaster) (fragment)
 C:Species: Drosophila melanogaster
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Feb-1997
 C:Accession: PS0443
 R:Adelman, J.P.; Shen, K.Z.; Kavanaugh, M.P.; Warren, R.A.; Wu, Y.N.; Lagrutta, A.; E
 Neuron 9, 209-216, 1992
 A:Title: Calcium-activated potassium channels expressed from cloned complementary DNA
 A:Reference number: JH0697; MUID:92360298; PMID:1497890
 A:Accession: PS0443
 A>Status: nucleic acid sequence not shown
 A:Molecule type: mRNA

A:Residues: 1-13 <ADE>
C:Comment: This potassium channel is activated by calcium.
C:Genetics:
A:Gene: FlyBase:slo
A:Cross-references: FlyBase:FBgn0003429
C:Keywords: alternative splicing; ion channel; potassium channel; transmembrane protein

Query Match 27.3%; Score 18; DB 2; Length 13;
Best Local Similarity 75.0%; Pred. No. 3e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 TKFD 12
:|:
Db 7 SKFD 10

RESULT 15
PT0077
proteochondroitin core protein - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 04-Sep-1998
C:Accession: PT0077
R;Marcum, J.A.; Thompson, M.A.
Biochem. Biophys. Res. Commun. 175, 706-712, 1991
A:Title: The amino-terminal region of a proteochondroitin core protein, secreted by aorta from human bone.
A:Reference number: PT0077; PMID:91207372; PMID:2018513
A:Accession: PT0077
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <MAR>
C:Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 27.3%; Score 18; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14
| | | |
Db 8 FWDFTL 14

Search completed: February 26, 2003, 14:58:06
Job time : 45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:55 ; Search time 10 Seconds
(without alignments)
58.067 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 633

Minimum DB seq length: 0
Maximum DB seq length: 14

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	40.9	14	1	LECB_PSOSC
2	19	28.8	13	1	ODPA_CANFA
3	18	27.3	10	1	UXA6_CHLTR
4	18	27.3	13	1	TEML_RANTE
5	18	27.3	14	1	DHSL_ANACY
6	17	25.8	10	1	LCMS_LEUMA
7	17	25.8	10	1	TKL4_LOCM1
8	17	25.8	10	1	TKN1_SCYCA
9	16	24.2	8	1	ACT_CARNA
10	16	24.2	10	1	FARP_MANSE
11	16	24.2	11	1	BRK_MEGFL
12	16	24.2	11	1	MLG_THETS
13	15	22.7	10	1	TRP9_LEUMA
14	15	22.7	12	1	V25K_WSSV
15	14	21.2	9	1	FAR3_CALVO
16	14	21.2	9	1	UHA2_HUMAN
17	14	21.2	10	1	FAR6_PANRE
18	14	21.2	10	1	MOSQ_CLYJA
19	14	21.2	12	1	UR2_POLSP
20	14	21.2	13	1	AU11_LITRA
21	14	21.2	13	1	AU12_LITRA
22	14	21.2	13	1	FIBA_CAVPO
23	14	21.2	13	1	NP2_LYMST
24	14	21.2	13	1	NP5_LYMST
25	13	19.7	4	1	PLRE_HIRME
26	13	19.7	4	1	FMRE_MACNI
27	13	19.7	7	1	FAR1_ACSU
28	13	19.7	7	1	FAR1_HELTI
29	13	19.7	7	1	FAR1_PROCL
30	13	19.7	7	1	FAR2_PROCL
31	13	19.7	7	1	FAR4_PANRE
32	13	19.7	7	1	FARB_CALVO
33	13	19.7	8	1	ALL4_CYDPO

34	13	19.7	8	1	FAR1_PANRE	P41872 panagrellus
35	13	19.7	8	1	FAR3_HOMAM	P41486 homarus ame
36	13	19.7	8	1	FAR4_HOMAM	P41487 homarus ame
37	13	19.7	8	1	FAR7_ACSU	P43171 ascaris suu
38	13	19.7	8	1	FAR8_CALVO	P41863 calliphora
39	13	19.7	9	1	FAR1_CALVO	P41856 calliphora
40	13	19.7	9	1	FAR2_CALVO	P41857 calliphora
41	13	19.7	9	1	FAR2_PANRE	P41873 panagrellus
42	13	19.7	9	1	FAR4_CALVO	P41859 calliphora
43	13	19.7	9	1	FAR5_ACSU	P43170 ascaris suu
44	13	19.7	9	1	FAR5_CALVO	P41860 calliphora
45	13	19.7	9	1	FAR6_CALVO	P41861 calliphora

ALIGNMENTS

RESULT 1
LECB_PSOSC
ID LECB_PSOSC STANDARD; PRT; 14 AA.
AC P22584;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Basic lectin B1 (Fragment).
OS Psophocarpus scandens.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
OC Psophocarpus.
OX NCBI_TaxID=3890;
RN [1]
RC SEQUENCE.
RP TISSUE=Seed;
RA Kott A.A.;
RT Isolation and characterization of the lectins from the seeds of
RT Psophocarpus scandens.;
RL Phytochemistry 27:2847-2855(1988).
CC -!- SUBUNIT: DIMER OF TWO IDENTICAL OR NEARLY IDENTICAL SUBUNITS OF
CC ABOUT 32000 APPARENT MW.
CC -!- PTM: CONTAINS ABOUT 9% CARBOHYDRATE. HETEROGENEITY IN CARBOHYDRATE
CC COMPOSITION MAY RESULT IN DIFFERENT ISOLECTIN FORMS.
CC -!- SIMILARITY: WITH P.TETRAGONOLOBUS BASIC LECTINS IN N-TERMINAL
CC SEQUENCE, AA COMPOSITION, ERYTHROCYTE AND SUGAR SPECIFICITIES.
DR PIR; PA0007; PA0007.
KW Lectin; Glycoprotein.
FT NON_TER 14
SQ SEQUENCE 14 AA; 1732 MW; D804CB43B487C549 CRC64;
Query Match 40.9%; Score 27; DB 1; Length 14;
Best Local Similarity 40.0%; Pred. No. 26;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 2 QSLSFXTKFK 11
:::| | | |
Db 1 ETISFNQF 10
RESULT 2
ODPA_CANFA STANDARD; PRT; 13 AA.
AC P49823;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pyruvate dehydrogenase E1 component alpha subunit, somatic form
DE (EC 1.2.4.1) (PDH-E1-A type I) (Fragment).
GN PDH1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]

```

RP SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RL dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
CC acetyldihydrolipoamide + CO(2).
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC -!- ENZYME REGULATION: E1 ACTIVITY IS REGULATED BY PHOSPHORYLATION
CC (INACTIVATION) AND DEPHOSPHORYLATION (ACTIVATION) OF THE ALPHA
CC SUBUNIT (BY SIMILARITY).
CC -!- SUBUNIT: TETRAMER OF TWO ALPHA AND TWO BETA SUBUNITS.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
DR HSC-2DPAGE; P49823; DOG.
KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
KW Mitochondrion; Phosphorylation.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1510 MW; C97EEBF844085B19 CRC64;

Query Match 28.8%; Score 19; DB 1; Length 13;
Best Local Similarity 44.4%; Pred. No. 7.8e+02;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 SPXFTKFDL 13
Db : : : :
5 TPEIKKXDL 13

RESULT 3
OXAG6_CHLTR
ID UXA6_CHLTR STANDARD; PRT; 10 AA.
AC P38007;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Unknown protein from 2D-page from elementary body (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE.
RC STRAIN=L2/434/Bu;
RA Biol L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA Comanducci M., Christensen G., Birkelund S., Vretou E., Ratti G.,
RA Pallini V.;
RL Submitted (SEP-1994) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 6.44, ITS MW IS: 38.6 kDa.
DR Siena-2DPAGE; P38007; -.
FT NON_TER 10 10
SQ SEQUENCE 10 AA; 1243 MW; DAD39A33304B5339 CRC64;

Query Match 27.3%; Score 18; DB 1; Length 10;
Best Local Similarity 33.3%; Pred. No. 9.2e+02;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 LSFXTKFD 12
Db : : : :
1 MNFKYIKD 9

RESULT 4
TEML_RANTE
ID TEML_RANTE STANDARD; PRT; 13 AA.
AC P57104;
DT 16-OCT-2001 (Rel. 40, Created)

```

```

DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Temporin L.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranolidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=97115050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Barra D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC GRAM-POSITIVE BACTERIA.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
KW Amphibian skin; Antibiotic; Amidation; Multigene family.
FT MOD_RES 13 13
SQ SEQUENCE 13 AA; 1641 MW; 9EBCB1FAFF7C325 CRC64;

Query Match 27.3%; Score 18; DB 1; Length 13;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 FTKF 11
Db : : : :
5 FSKF 8

RESULT 5
DHSL_ANACY
ID DHSL_ANACY STANDARD; PRT; 14 AA.
AC P17874;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Soluble hydrogenase 50 kDa subunit (EC 1.12.-.-) (Fragment).
OS Anabaena cylindrica.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1165;
RN [1]
RP SEQUENCE.
RX MEDLINE=90126821; PubMed=2129525;
RA Ewart G.D., Reed K.C., Smith G.D.;
RT "Soluble hydrogenase of Anabaena cylindrica. Cloning and sequencing
RT of a potential gene encoding the tritium exchange subunit.";
RL Eur. J. Biochem. 187:215-223(1990).
CC -!- FUNCTION: SOLUBLE HYDROGENASE CATALYZES BOTH PRODUCTION AND
CC CONSUMPTION OF HYDROGEN FROM SUITABLE ARTIFICIAL ELECTRON DONORS
CC OR ACCEPTORS. THIS SUBUNIT (50 kDa) IS REQUIRED FOR HYDROGEN
CC PRODUCTION WITH REDUCED METHYL-VIOLOGEN.
CC -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
KW Oxidoreductase.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1551 MW; 9254DAFB141CFF2A CRC64;

Query Match 27.3%; Score 18; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 10 KFDLD 14
Db : : : :
3 EFDVD 7

RESULT 6

```

LCMS_LEUMA
ID LCMS_LEUMA STANDARD; PRT; 10 AA.
AC P21144; P41197;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Leucomyosuppressin (LMS) (Lem-MS).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Head;
RX Holman G.M., Cook B.J., Nachman R.J.;
RT "Isolation, primary structure and synthesis of leucomyosuppressin,
RT an insect neuropeptide that inhibits spontaneous contractions of the
RT cockroach hindgut.";
RL Comp. Biochem. Physiol. 85C:329-333(1986).
CC -1- FUNCTION: INHIBITS THE SPONTANEOUS CONTRACTIONS OF COCKROACH
CC PROTODEUM (HINDGUT).
KW Neuropeptide; Amidation.
FT MOD_RES 1 1
FT MOD_RES 10 10
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1275 MW; D3C45229D2C1EAB2 CRC64;
Query Match 25.8%; Score 17; DB 1; Length 10;
Best Local Similarity 30.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 QSLSFXTKFK 11
I : I : I
Db 1 QDVDFVFLRF 10

RESULT 7
TKL4_LOCMI
ID TKL4_LOCMI STANDARD; PRT; 10 AA.
AC P30250;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Locustatachykinin IV (TK-IV).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=91219696; PubMed=2132575;
RA Schoofs L., Holman G.M., Hayes T.K., Kochansky J.P., Nachman R.J.,
RA de Loof A.;
RT "Locustatachykinin III and IV: two additional insect neuropeptides
RT with homology to peptides of the vertebrate tachykinin family.";
RL Regul. Pept. 31:199-212(1990).
CC -1- FUNCTION: MFOACTIVE PEPTIDE. STIMULATES THE CONTRACTION OF THE
CC OVIDUCT AND FOREGUT.
CC OVIDUCT AND FOREGUT.
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
DR PIR; B60073; ECLQ4M.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1040 MW; 9E52CD71E9C87735 CRC64;
Query Match 25.8%; Score 17; DB 1; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 AQSLSF 6
I : I : I
Db 1 APSLSGF 6

RESULT 8
TKNI_SCYCA
ID TKNI_SCYCA STANDARD; PRT; 10 AA.
AC P08608;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Scyllorhinus I.
OS Scyllorhinus canicula (Spotted dogfish) (Spotted catshark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphi; Galeoidea; Carchariniiformes;
OC Scyllorhinidae; Scyllorhinus.
OX NCBI_TaxID=7830;
RN [1]
RP SEQUENCE.
RX MEDLINE=86192829; PubMed=2422058;
RA Conlon J.M., Deacon C.F., O'Toole L., Thim L.;
RT "Scyllorhinin I and II: two novel tachykinins from dogfish gut.";
RL FEBS Lett. 200:111-116(1986).
RN [2]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=93292508; PubMed=7685693;
RA Waugh D., Wang Y., Hazon N., Balmont R.J., Conlon J.M.;
RT "Primary structures and biological activities of substance-P-related
RT peptides from the brain of the dogfish, Scyllorhinus canicula.";
RL Eur. J. Biochem. 214:469-474(1993).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR PIR; A24867; A24867.
DR PIR; S33301; S33301.
DR InterPro; IPR002040; Tachykinin.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1219 MW; D0602D6B59C33AA9 CRC64;
Query Match 25.8%; Score 17; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 KFD 12
I : I : I
Db 2 KFD 4

RESULT 9
ACT_CARMA
ID ACT_CARMA STANDARD; PRT; 8 AA.
AC P80709;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Actin (Fragment).
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Pleocyemata;
OC Brachyura; Eubrachyura; Portunioidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RA Lachaise F., Somme G., Carpentier G., Granjeon E., Webster S.,
RA Baghdassarian D.;
RT "A transaldolase. An enzyme implicated in crab steroidogenesis.";
RL Endocrine 5:23-32(1996).
CC -1- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
CC IN ALL EUKARYOTIC CELLS.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
 CC 6.8, ITS MW IS: 46 kDa.
 CC -!- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
 DR InterPro: IPR004001; Actin.
 DR InterPro: IPR004000; Actin_like.
 DR PROSITE: PS00406; ACTINS_1; PARTIAL.
 DR PROSITE: PS00432; ACTINS_2; PARTIAL.
 DR PROSITE: PS01132; ACTINS_ACT_LIKE; PARTIAL.
 KW Structural protein.
 FT NON_TER 1
 FT NON_TER 8
 SQ SEQUENCE 8 AA; 976 MW; 1424005AB2CAAB3 CRC64;
 Query Match 24.2%; Score 16; DB 1; Length 8;
 Best Local Similarity 60.0%; Pred. No. 1.1e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KFDLD 14
 I I I
 Db 1 KCDVD 5
 RESULT 10
 FARP_MANSE STANDARD; PRT; 10 AA.
 AC P18523;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE FMRamide-like neuropeptide.
 OS Manduca sexta (tobacco hawkmoth) (Tobacco hornworm).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 CC Ditrysia; Spingioidea; Sphingidae; Sphinginae; Manduca.
 OX NCBI_TaxID=7130;
 RN [1]
 RP SEQUENCE
 RX MEDLINE=91045350; PubMed=2235684;
 RA Kingan T.G., Teplow D.B., Phillips J.M., Riehm J.P., Rao K.R.,
 RA Hildebrand J.G., Homborg U., Kammer A.E., Jardine I., Griffin P.R.,
 RA Hunt D.F.;
 RT "A new peptide in the FMRamide family isolated from the CNS of the
 RT hawkmoth, Manduca sexta";
 RL Peptides 11:849-856(1990).
 CC -!- FUNCTION: INCREASES THE FORCE OF NEURALLY EVOKED CONTRACTIONS IN
 CC THE MAJOR POWER-PRODUCING FLIGHT MUSCLES, THE DORSAL LONGITUDINAL
 CC MUSCLES AND SO IS LIKELY TO PLAY A ROLE IN SUSTAINING OR PROMOTING
 CC FLIGHT BEHAVIOR PATTERNS.
 CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
 CC FAMILY.
 DR PIR: A43977; A43977.
 KW Amidation; Neuropeptide.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1247 MW; D3C45229D5B1F2D2 CRC64;
 Query Match 24.2%; Score 16; DB 1; Length 10;
 Best Local Similarity 30.0%; Pred. No. 2.2e+03;
 Matches 3; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 2 QSLSFXTK 11
 I I I I I
 Db 1 QDVVHSFURF 10
 RESULT 11
 BRK_MEGFL STANDARD; PRT; 11 AA.
 AC P12797;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Megascoliakinin ([Thr6]bradykinin-Lys-Ala) [Contains: Bradykinin-like
 DE peptide ([Thr6]bradykinin)].
 OS Megascollia flavifrons (Garden dagger wasp) (Solitary wasp).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
 CC Aculeata; Scolioidea; Scollidae; Megascollia.
 OX NCBI_TaxID=7437;
 RN [1]
 RP SEQUENCE
 RC TISSUE=Venom;
 RX MEDLINE=87293024; PubMed=3617088;
 RA Yasuhara T., Mantel P., Nakajima T., Piek T.;
 RT "Two kinins isolated from an extract of the venom reservoirs of the
 RT solitary wasp Megascollia flavifrons.";
 RL Toxicon 25:527-535(1987).
 RN [2]
 RP SEQUENCE
 RC TISSUE=Venom;
 RA Nakajima T., Piek T., Yasuhara T., Mantel P.;
 RT "Two kinins isolated from the venom of Megascollia flavifrons.";
 RL Toxicon 26:34-34(1988).
 CC -!- FUNCTION: BOTH PROTEINS HAVE BRADYKININ-LIKE, ALTHOUGH LOWER
 CC ACTIVITIES (E.G. SMOOTH MUSCLE CONTRACTION).
 CC -!- SUBCELLULAR LOCATION: SECRETED; WASP VENOM RESERVOIRS.
 CC -!- SIMILARITY: BELONGS TO THE BRADYKININ FAMILY.
 DR PIR: A26744; A26744.
 DR PIR: B26744; B26744.
 KW Bradykinin; Vasodilator; Venom.
 FT PEPTIDE 1 11
 FT PEPTIDE 1 9 BRADYKININ-LIKE PEPTIDE.
 SQ SEQUENCE 11 AA; 1273 MW; 33867393D771A9C8 CRC64;
 Query Match 24.2%; Score 16; DB 1; Length 11;
 Best Local Similarity 75.0%; Pred. No. 2.4e+03;
 Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 8 FTKF 11
 I I I
 Db 5 FTFP 8
 RESULT 12
 MLG_THETS STANDARD; PRT; 11 AA.
 AC P41889;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanotropin gamma (Gamma-melanocyte stimulating hormone) (Gamma-MSH).
 OS Theromyzon tessulatum (Leech).
 CC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
 CC Rhynchobdellida; Glossiphoniidae; Theromyzon.
 OX NCBI_TaxID=13286;
 RN [1]
 RP SEQUENCE
 RC TISSUE=Brain;
 RX MEDLINE=94298944; PubMed=8026574;
 RA Salzet M., Watez C., Bulet P., Malecha J.;
 RT "Isolation and structural characterization of a novel peptide related
 RT to gamma-melanocyte stimulating hormone from the brain of the leech
 RT Theromyzon tessulatum";
 RL FEBS Lett. 348:102-106(1994).
 CC -!- SIMILARITY: BELONGS TO THE POMC FAMILY.
 KW Hormone; Amidation.
 FT MOD_RES 11 11 AMIDATION.
 SQ SEQUENCE 11 AA; 1486 MW; 2DB8FACE6409C1E8 CRC64;
 Query Match 24.2%; Score 16; DB 1; Length 11;
 Best Local Similarity 50.0%; Pred. No. 2.4e+03;
 Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 6 FXFTKF 11

Db 6 FRWDKF 11

RESULT 13

TRP9_LEUMA
ID TRP9_LEUMA STANDARD; PRT; 10 AA.
AC P81741;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Tachykinin-related peptide 9 (LemRP 9).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
OX NCBI_TaxID=6988;
RN [1]

SEQUENCE, AND MASS SPECTROMETRY.

RP TISSUE=Brain;
RC MEDLINE=97269266; PubMed=9114447;
RX Muren J.E., Naessel D.R.;
RA "Seven tachykinin-related peptides isolated from the brain of the
RT Madeira cockroach; evidence for tissue-specific expression of
RT isoforms.";
RL Peptides 18:7-15(1997).
CC -1- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- MASS SPECTROMETRY: MW=1081.5; METHOD=MALDI.
CC -1- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1081 MW; 9E469D66D9C87685 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 10;
Best Local Similarity 50.0%; Pred. No. 3.4e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSF 6
| | | |
Db 1 APSMGF 6

RESULT 14

V25K_WSSV
ID V25K_WSSV STANDARD; PRT; 12 AA.
AC P82004;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 25 kDa structural polyprotein (Fragment).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE.
RC STRAIN=South Carolina;
RX MEDLINE=20214217; PubMed=10752552;
RA Wang Q., Poulos B.T., Lightner D.V.;
RT "Protein analysis of geographic isolates of shrimp white spot syndrome
RT virus.";
RL Arch. Virol. 145:263-274(2000).
CC -1- FUNCTION: STRUCTURAL COMPONENT OF THE VIRION.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1283 MW; C5409AD9ECB731A9 CRC64;

Query Match 22.7%; Score 15; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 4.1e+03;
Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 LSFXTF 9
| | | |

Db 3 LSFTLS 8

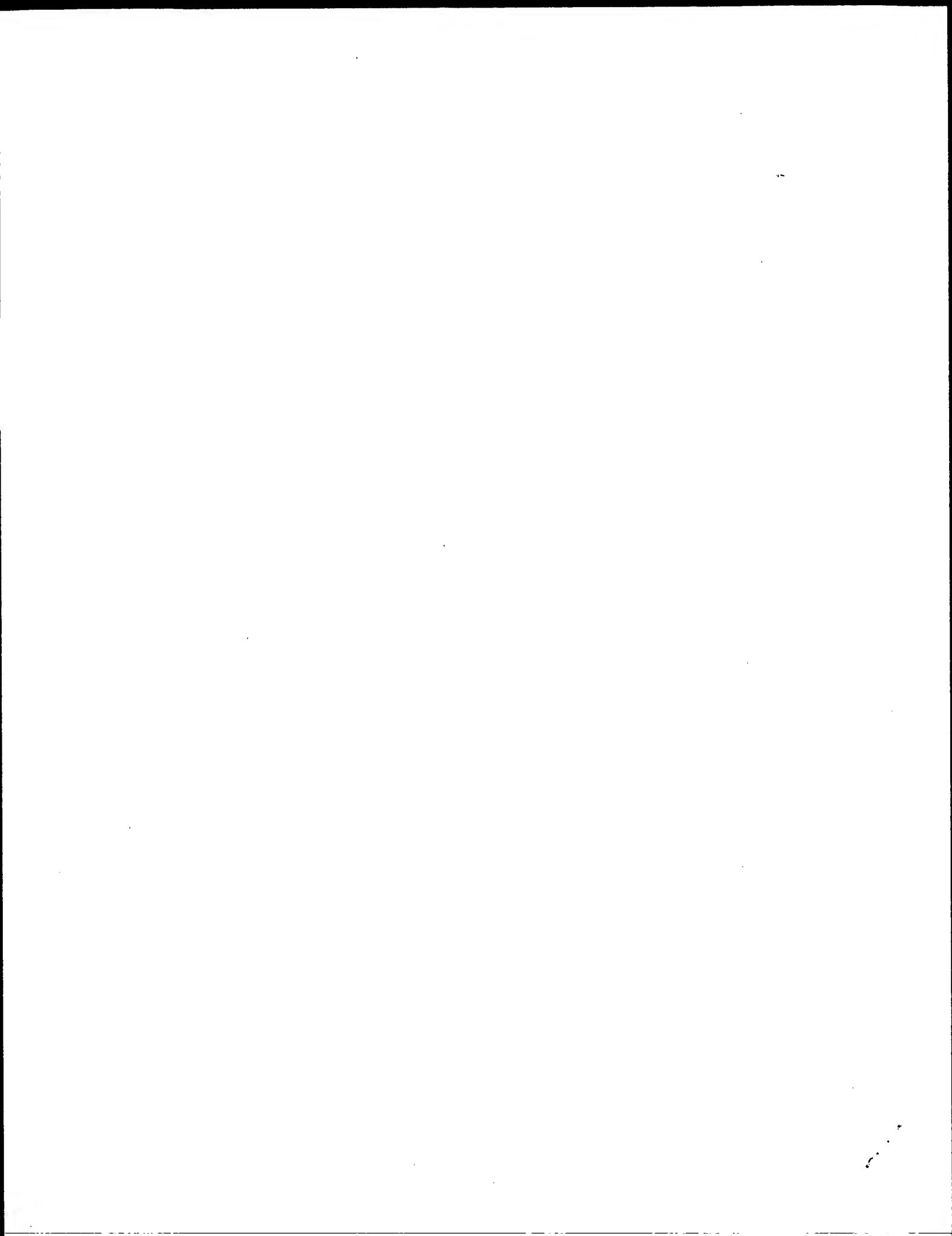
RESULT 15

FAR3_CALVO
ID FAR3_CALVO STANDARD; PRT; 9 AA.
AC P41858;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CallifMRamide 3.
OS Calliphora vomitoria (Blue blowfly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Oestroidea; Calliphoridae; Calliphora.
OX NCBI_TaxID=27454;
RN [1]
RP SEQUENCE.
RC TISSUE=Thoracic ganglion;
RX MEDLINE=92196111; PubMed=1549595;
RA Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA Rehfeld J.F., Thorpe A.;
RT "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT neuropeptides (designated callifMRFamides) from the blowfly
RT Calliphora vomitoria.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC -1- FUNCTION: ABLE TO INDUCE FLUID SECRETION FROM THE ISOLATED
CC SALIVARY GLAND OF CALLIPHORA.
CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
CC PIR; C41978; C41978.
KW Neuropeptide; Amidation.
FT MOD_RES 9 9
SQ SEQUENCE 9 AA; 1114 MW; 2F0B0699CAB6C5A7 CRC64;

Query Match 21.2%; Score 14; DB 1; Length 9;
Best Local Similarity 44.4%; Pred. No. 1.1e+05;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SLSFXFTKF 11
| | | | |
Db 1 SPSQDFMRF 9

Search completed: February 26, 2003, 14:56:39
Job time : 10 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:54:20 ; Search time 28 Seconds

(without alignments)
103.024 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 2683

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	29	43.9	8	3 P87225	P87225 saccharomyc
2	27	40.9	14	4 Q96KF9	Q96KF9 homo sapien
3	24	36.4	9	4 Q14277	Q14277 homo sapien
4	19	28.8	13	6 Q9TU76	Q9TU76 ovis aries
5	19	28.8	14	2 Q9R518	Q9R518 vibrio algi
6	18	27.3	9	2 Q8RKC6	Q8RKC6 erwinia chr
7	18	27.3	14	15 Q9PYL3	Q9PYL3 human t-cel
8	17	25.8	10	11 Q8VHM9	Q8VHM9 mus musculus
9	17	25.8	10	13 Q73588	Q73588 gallus gall
10	17	25.8	12	4 Q96PK0	Q96PK0 homo sapien
11	17	25.8	13	4 Q9UCT1	Q9UCT1 homo sapien
12	17	25.8	13	11 Q63047	Q63047 rattus norv
13	17	25.8	14	2 Q9LCS1	Q9LCS1 bacillus su
14	17	25.8	14	2 Q8VQ14	Q8VQ14 micrococci
15	17	25.8	14	4 Q96Q62	Q96Q62 homo sapien
16	17	25.8	14	10 P82435	P82435 nicotiana t

17	17	25.8	14	11 Q9Z1H4	Q9Z1H4 mus musculus
18	17	25.8	14	12 Q67112	Q67112 influenzavi
19	16	24.2	7	2 Q07354	Q07354 synechococc
20	16	24.2	9	8 Q8WGE6	Q8WGE6 procambary
21	16	24.2	9	13 P83059	P83059 bombina ori
22	16	24.2	9	13 P83058	P83058 bombina var
23	16	24.2	9	13 P83057	P83057 bombina var
24	16	24.2	9	13 P83056	P83056 bombina var
25	16	24.2	10	4 Q60912	Q60912 homo sapien
26	16	24.2	10	6 Q9TRC1	Q9TRC1 bos taurus
27	16	24.2	10	10 Q9S936	Q9S936 beta vulgar
28	16	24.2	11	2 Q9AIZ8	Q9AIZ8 carsonella
29	16	24.2	11	2 Q9AIV6	Q9AIV6 carsonella
30	16	24.2	11	7 Q78119	Q78119 oreochromis
31	16	24.2	12	11 Q925V7	Q925V7 mus musculus
32	16	24.2	13	2 Q82835	Q82835 synechococc
33	16	24.2	14	2 Q9R8L2	Q9R8L2 chlamydia t
34	16	24.2	14	2 Q9R8J2	Q9R8J2 chlamydia t
35	16	24.2	14	2 Q9R8J0	Q9R8J0 chlamydia t
36	16	24.2	14	2 Q9R8I8	Q9R8I8 chlamydia t
37	16	24.2	14	2 Q9R8I6	Q9R8I6 chlamydia t
38	16	24.2	14	2 Q9R8I4	Q9R8I4 chlamydia t
39	16	24.2	14	2 Q9R8I2	Q9R8I2 chlamydia t
40	16	24.2	14	2 Q9R8I0	Q9R8I0 chlamydia t
41	16	24.2	14	2 Q9R8H8	Q9R8H8 chlamydia t
42	16	24.2	14	2 Q9R8H6	Q9R8H6 chlamydia t
43	16	24.2	14	2 Q9R8H4	Q9R8H4 chlamydia t
44	16	24.2	14	2 Q9R8H2	Q9R8H2 chlamydia t
45	16	24.2	14	2 Q9R8H0	Q9R8H0 chlamydia t

ALIGNMENTS

RESULT 1

P87225
ID P87225 PRELIMINARY; PRT; 8 AA.
AC P87225;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE GIN11 protein (Fragment).
GN GIN11 OR YLL065W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Wedler H., Wedler E., Scharfe M., Wambutt R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 273159; CAA97518.2; -.
DR SGD; S0003988; GIN11.
FT NON_TER 1
SQ SEQUENCE 8 AA; 1019 MW; 4E21A9C449D5B73B CRC64;

Query Match 43.9%; Score 29; DB 3; Length 8;
Best Local Similarity 85.7%; Pred. No. 6.7e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 LSFXTFK 10
||| |||
DB 2 LSFNFTK 8

RESULT 2

Q96KF9
ID Q96KF9 PRELIMINARY; PRT; 14 AA.
AC Q96KF9;

```

DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Putative magnesium transporter (Fragment).
GN MRS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE 14 AA; 1537 MW; C6344FF1E984AFB2 CRC64;
RC TISSUE=BL00D;
RX MEDLINE=21295035; PubMed=11401429;
RA Zsuzka G., Gregan J., Schweyen R.J.;
RT "The human mitochondrial MRS2 protein functionally substitutes for its
RL yeast homologue, a candidate magnesium transporter.";
DR EMBL: AF293077; AAK38616.1; -.
FT NON_TER 1 14
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1537 MW; C6344FF1E984AFB2 CRC64;

Query Match 40.9%; Score 27; DB 4; Length 14;
Best Local Similarity 71.4%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 6 FXTKFD 12
Db 5 FVTKFD 11

RESULT 3
Q14277 PRELIMINARY; PRT; 9 AA.
AC Q14277;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE RET protein short form (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94071887; PubMed=7902707;
RA Ceccherini I., Boccardi R., Luo Y., Pasini B., Hofstra R.,
RA Takahashi M., Romeo G.;
RT "Exon structure and flanking intronic sequences of the human RET
RL proto-oncogene.";
RN Biochem. Biophys. Res. Commun. 196:1288-1295(1993).
RP SEQUENCE FROM N.A.
RX MEDLINE=94366753; PubMed=8084609;
RA Ceccherini I., Hofstra R., Yin L., Stulp R., Barone V., Stelwagen T.,
RA Boccardi R., Nijveen H., Bolino A., Seri M., Ronchetto P., Pasini B.,
RA Bozzano M., Buys C., Romeo G.;
RT "DNA polymorphisms and conditions for SSCP analysis of the 20 exons of
RL the ret proto-oncogene.";
RN Oncogene 9:3025-3029(1994).
DR EMBL: U11532; AAC50102.1; -.
FT NON_TER 1 1
FT NON_TER 1 1
SQ SEQUENCE 9 AA; 1134 MW; 034E11A9CDD1F5B0 CRC64;

Query Match 36.4%; Score 24; DB 4; Length 9;
Best Local Similarity 50.0%; Pred. No. 6.7e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LSFYTKF 11
Db 2 ISHAFTRF 9

```

```

RESULT 4
Q9TU76 PRELIMINARY; PRT; 13 AA.
AC Q9TU76;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE Interleukin 1 alpha (Fragment).
GN IL1A.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99397042; PubMed=10467711;
RA Maddox J.F., Hawken R.J., Matthew P., Davies K.P.;
RT "Single strand conformational polymorphisms (SSCPs) in the ovine IL1A
RL and IL6 genes.";
DR Anim. Genet. 30:317-318(1999).
FT NON_TER 1 1
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1520 MW; C26BC6198305DB5D CRC64;

Query Match 28.8%; Score 19; DB 6; Length 13;
Best Local Similarity 75.0%; Pred. No. 5e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDL 13
Db 1 KFDL 4

RESULT 5
Q9R518 PRELIMINARY; PRT; 14 AA.
AC Q9R518;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE L-2,4-diaminobutyrate decarboxylase (Fragment).
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=663;
RN [1]
RP SEQUENCE.
RX MEDLINE=92381494; PubMed=1512577;
RA Yamamoto S., Tsuzaki Y., Tougou K., Shinoda S.;
RT "Purification and characterization of L-2,4-diaminobutyrate
RL decarboxylase from Acinetobacter calcoaceticus.";
RN J. Gen. Microbiol. 138:1461-1465(1992).
SQ SEQUENCE 14 AA; 1643 MW; 9F1B13DD35168ABA CRC64;

Query Match 28.8%; Score 19; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 5.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 TKFDLD 14
Db 2 TAFEVD 7

RESULT 6
Q8RKC6 PRELIMINARY; PRT; 9 AA.
AC Q8RKC6;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Expr protein (Fragment).
GN EXPR.

```


OS Erwinia chrysanthemi.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Pectobacterium.
 OX NCBI_TaxID=556;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=3937;
 RA Reverchon S.;
 RT "Identification of a lysA-like gene required for virulence factors
 synthesis in Erwinia chrysanthemi.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ438189; CAD27339.1; -
 FT NON_TER 9
 SQ SEQUENCE 9 AA; 999 MW; 9A8BC455B9D5B045 CRC64;
 Query Match 27.3%; Score 18; DB 2; Length 9;
 Best Local Similarity 33.3%; Pred. No. 6.7e+05;
 Matches 3; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 QY 4 LSFXTKFD 12
 Db : | : |
 1 MSISFSNVD 9
 RESULT 7
 Q9PYL3 PRELIMINARY; PRT; 14 AA.
 AC Q9PYL3;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TReMBLrel. 13, Last annotation update)
 DE Transcriptional transactivator Tax (Fragment).
 GN TAX.
 OS Human T-cell leukemia virus type II (HTLV-II).
 OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.
 OX NCBI_TaxID=11909;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H2BC1;
 RX MEDLINE=20014165; PubMed=10548125;
 RA Peters A.A., Oger J.J., Coulthart M.B., Waters D.J., Cummings H.J.,
 RA Dekaban G.A.;
 RT "An apparent case of human T-cell lymphotropic virus type II (HTLV-II)-
 associated neurological disease: a clinical, molecular, and
 phylogenetic characterisation.";
 RL J. Clin. Virol. 14:37-50(1999).
 DR EMBL; AF115495; AAF15550.1; -
 FT NON_TER 1
 SQ SEQUENCE 14 AA; 1567 MW; 7353D4CA8CC6041B CRC64;
 Query Match 27.3%; Score 18; DB 15; Length 14;
 Best Local Similarity 40.0%; Pred. No. 8.5e+03;
 Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 5 SFXFTKFDLD 14
 Db : | : |
 1 SILENKBEAD 10
 RESULT 8
 Q8VHM9 PRELIMINARY; PRT; 10 AA.
 AC Q8VHM9;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Interferon receptor 2a' (Fragment).
 GN IFNAR2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Hardy M.P., Hertzog P.J., Owczarek C.M.;
 RT "The genomic structure and expression patterns of the gene encoding
 the second chain of the murine interleukin 10 receptor, IL-10R2.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Hardy M.P., Hertzog P.J., Owczarek C.M.;
 RT "The organization, transcriptional regulation and chromosomal
 localization of the locus encoding the gene for the murine type I
 interferon receptor, Ifnar2.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF440786; AAL40944.1; -
 KW Receptor.
 FT NON_TER 1
 SQ SEQUENCE 10 AA; 1139 MW; 322A36A72449C737 CRC64;
 Query Match 25.8%; Score 17; DB 11; Length 10;
 Best Local Similarity 75.0%; Pred. No. 9.5e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 11 FDL 14
 Db : | : |
 5 FNLD 8
 RESULT 9
 O73588 PRELIMINARY; PRT; 10 AA.
 AC O73588;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Engrailed-3 (Fragment).
 GN EN-3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WHITE LEHORN; TISSUE=EMBRYO;
 RX MEDLINE=98141813; PubMed=9473273;
 RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;
 RT "Multiplex display polymerase chain reaction amplifies and resolves
 related sequences sharing a single moderately conserved domain.";
 RL Anal. Biochem. 256:158-168(1998).
 DR EMBL; U26148; AAC06186.1; -
 FT NON_TER 1
 SQ SEQUENCE 10 AA; 1118 MW; 73C0BE14735B72B CRC64;
 Query Match 25.8%; Score 17; DB 13; Length 10;
 Best Local Similarity 80.0%; Pred. No. 9.5e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 AQSL 5
 Db : | : |
 1 AQELS 5
 RESULT 10
 Q96PK0 PRELIMINARY; PRT; 12 AA.
 AC Q96PK0;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE COPG2 (Fragment).
 GN COPG2.

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bonora E., Bacchelli E., Levy E.R., Blasi F., Marlow A., Monaco A.P.,
RA Maestrini E.;
RT "Mutation screening and imprinting analysis of four candidate genes
RT for autism in the 7q32 region.";
RL Mol. Psychiatry 0:0-0(2001).
DR EMBL; AF324497; AAL17778.1; -.
FT NON_TER 12
SQ SEQUENCE 12 AA; 1498 MW; C344E6689A333AA9 CRC64;

Query Match 25.8%; Score 17; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFD 12
Db 4 KFD 6

RESULT 11
ID Q9UCT1 PRELIMINARY; PRT; 13 AA.
AC Q9UCT1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet glycoprotein GPIC (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=92061944; PubMed=1953640;
RA Catimel B., Parmentier S., Leung L.L., McGregor J.L.;
RT "Separation of important new platelet glycoproteins (GP1a, GPIC,
RT GP1c*, GP1a and GMP-140) by f.p.l.c. Characterization by monoclonal
RT antibodies and gas-phase sequencing.";
RL Biochem. J. 279:419-425(1991).
FT NON_TER 1
FT NON_TER 13
SQ SEQUENCE 13 AA; 1620 MW; 79740A62649AB04 CRC64;

Query Match 25.8%; Score 17; DB 4; Length 13;
Best Local Similarity 75.0%; Pred. No. 1.2e+04;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FDL 14
Db 1 FDL 4

RESULT 12
ID Q63047 PRELIMINARY; PRT; 13 AA.
AC Q63047;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AMP deaminase (Fragment).
GN AMPD1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=SOLEUS MUSCLE;

```

```

RX MEDLINE=90377216; PubMed=2398891;
RA Mineo I., Clarke P.R.H., Sabina R.L., Holmes E.W.;
RT "A novel pathway for alternative splicing: Identification of an RNA
RT intermediate that generates an alternative 5' splice donor site not
RT present in the primary transcript of AMPD1.";
RL Mol. Cell. Biol. 10:5271-5278(1990).
DR EMBL; M58689; AAA40727.1; -.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1524 MW; 526C5A93EF6201A7 CRC64;

Query Match 25.8%; Score 17; DB 11; Length 13;
Best Local Similarity 42.9%; Pred. No. 1.2e+04;
Matches 3; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 FXFTKFD 12
Db 4 FKLTEID 10

RESULT 13
ID Q9LCSI PRELIMINARY; PRT; 14 AA.
AC Q9LCSI;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Levan sucrose (Fragment).
GN SACB.
OS Bacillus subtilis.
OG Plasmid p1257.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Maldonado R., Casadesus J.;
RT "Identification of IS210 in Azotobacter vinelandii: a novel,
RT functional insertion element member of the IS5 family.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249381; CAB76429.1; -.
KW Plasmid.
FT NON_TER 1
FT NON_TER 14
SQ SEQUENCE 14 AA; 1623 MW; 8DC9108BA1B18745 CRC64;

Query Match 25.8%; Score 17; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFD 12
Db 1 KFD 3

RESULT 14
ID Q8VQ14 PRELIMINARY; PRT; 14 AA.
AC Q8VQ14;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE ErMML leader peptide.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OG Plasmid pMEC2.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococaceae; Micrococcus.
OX NCBI_TaxID=1270;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAW843;
RA Liebl W., Kloos W.E., Ludwig W.;
RT "Plasmid-borne macrolide-lincosamide-streptogramin B (MLS) resistance
RT in Micrococcus luteus.";

```

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF462611; AAL68826.1; -
 KW Plasmid.

SQ SEQUENCE 14 AA; 1625 MW; C6BF8E5F5CD58BCD CRC64;

Query Match 25.8%; Score 17; DB 2; Length 14;
 Best Local Similarity 44.4%; Pred. No. 1.3e+04;
 Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 SLSFXFTKF 11

Db 3 SPSTAVTRF 11

RESULT 15

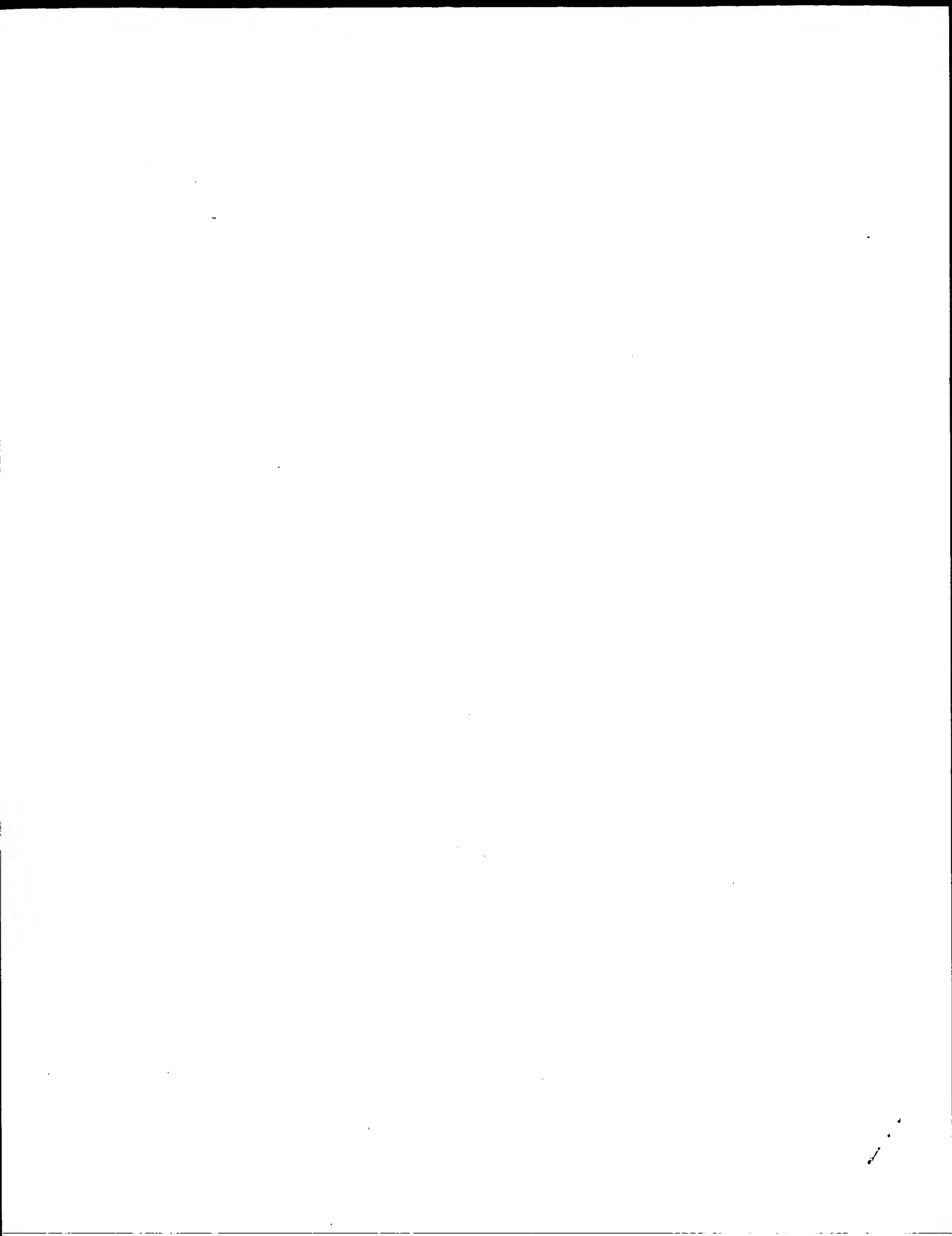
ID Q96Q62 PRELIMINARY; PRT; 14 AA.
 AC Q96Q62;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
 DE Mitochondrial ribosomal protein S11 (Fragment).
 GN MRPS11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21429115; PubMed=11543634;
 RA Kenmochi N., Suzuki T., Uechi T., Magoori M., Kuniba M., Higa S.,
 RA Watanabe K., Tanaka T.;
 RT "The human mitochondrial ribosomal protein genes: Mapping of 54 genes
 to the chromosomes and implications for human disorders.";
 RL Genomics 77:65-70(2001).
 DR EMBL; AB051349; BAB54939.2; -
 KW Ribosomal protein.
 FT NON_TER 1
 FT NON_TER 14
 SQ SEQUENCE 14 AA; 1585 MW; C07121F2234438DD CRC64;

Query Match 25.8%; Score 17; DB 4; Length 14;
 Best Local Similarity 40.0%; Pred. No. 1.3e+04;
 Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QSLFXFTKF 11

Db 5 QNAAPSHTKF 14

Search completed: February 26, 2003, 14:57:15
 Job time : 30 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:51:20 ; Search time 34 seconds
(without alignments)
54.868 Million cell updates/sec

Title: us-09-476-485a-31

Perfect score: 66

Sequence: 1 AOSLSFXFTKFDLD 14

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 228063

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

```

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	65	98.5	14	AAW61498
2	64	97.0	14	AAW67818
3	64	97.0	14	AAW62895
4	55	83.3	12	AAW61497
5	47	71.2	14	AAW62896
6	25	37.9	8	AAW62906
7	25	37.9	8	AAW63026
8	25	37.9	9	AAW46829
9	25	37.9	9	AAW62907
10	25	37.9	9	AAW63027

11	25	37.9	9	22	AAW98514
12	25	37.9	9	22	AAW98560
13	25	37.9	10	21	AAW78322
14	25	37.9	10	21	AAW62908
15	25	37.9	10	21	AAW63028
16	25	37.9	11	21	AAW12876
17	25	37.9	12	18	AAW27590
18	25	37.9	12	18	AAW39757
19	25	37.9	13	22	AAW05005
20	25	37.9	13	23	AAW86065
21	24	36.4	7	16	AAW71396
22	24	36.4	7	19	AAW38291
23	24	36.4	7	22	AAW37816
24	24	36.4	9	21	AAW39807
25	24	36.4	9	22	AAW10843
26	24	36.4	10	15	AAW49439
27	24	36.4	10	19	AAW76007
28	24	36.4	10	22	ABB56106
29	24	36.4	10	22	AAU28788
30	24	36.4	10	22	AAU26436
31	24	36.4	10	22	AAW61365
32	24	36.4	10	23	ABG67865
33	24	36.4	12	11	AAW06818
34	24	36.4	12	13	AAW20117
35	24	36.4	12	15	AAW49477
36	24	36.4	12	15	AAW49390
37	24	36.4	12	15	AAW49401
38	24	36.4	12	16	AAW69608
39	24	36.4	12	20	AAW25343
40	24	36.4	13	11	AAW06819
41	24	36.4	13	13	AAW20118
42	24	36.4	13	16	AAW69609
43	24	36.4	13	19	AAW77284
44	24	36.4	13	19	AAW44131
45	24	36.4	14	15	AAW49400

ALIGNMENTS

RESULT 1

AAW61498
ID AAW61498 standard; peptide; 14 AA.

AC AAW61498;

XX AAW61498;

XX AAW61498;

DT 19-OCT-1998 (first entry)

DE Pylartin protein, peptide chain beta (ii).

DE Pylartin protein, peptide chain beta (ii).

XX Pylartin protein; progenitor cell; haematopoietic system; cancer;

XX Pylartin protein; progenitor cell; haematopoietic system; cancer;

KW engraftation; haematologic disease; sickle cell anaemia; thalassemia.

XX engraftation; haematologic disease; sickle cell anaemia; thalassemia.

OS Leguminosae.

XX Leguminosae.

PN WO9825457-A1.

XX WO9825457-A1.

PD 18-JUN-1998.

XX 18-JUN-1998.

PF 09-DEC-1997; 97WO-US22486.

XX 09-DEC-1997; 97WO-US22486.

PR 28-MAR-1997; 97US-0825369.

PR 09-DEC-1996; 96US-0762537.

XX (IMCL-) IMCLONE SYSTEMS INC.

XX Moore JG;

XX Moore JG;

DR WPI; 1998-348161/30.

XX WPI; 1998-348161/30.

XX WPI; 1998-348161/30.

PT Pylartin proteins which preserve progenitor cells - useful for

PT haematopoietic therapies in cancer treatment or for treating

PT haematopoietic diseases

PT haematopoietic diseases

PT haematopoietic diseases

XX Claim 4; Page 28; 46pp; English.

XX The peptides AAW61497-W61502 can be used to form pylartin proteins which

CC preserve progenitor cells (pc). The proteins can be used for protecting

CC the integrity of the haematopoietic processes in vivo and as adjuncts in

CC therapeutic treatments related to cancer and other diseases which can

CC otherwise adversely impact upon the haematopoietic system. Since the

CC proteins bind specifically to primitive PCs, they can also be used for

CC the identification and localisation of PCs. The methods can be used for

CC e.g. expanding PC populations ex vivo to increase chances of

CC engraftation, improving conditions for transporting and storing PCs and

CC for removing a fundamental barrier thereby enabling gene therapy to

CC treat and cure a broad range of life-threatening haematologic diseases

CC such as sickle cell anaemia and thalassemia.

XX

SQ Sequence 14 AA;

Query Match 98.5%; Score 65; DB 19; Length 14;

Best Local Similarity 92.9%; Pred. No. 9.1e-06;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFSFTKFDLD 14

RESULT 2

AARG7818

ID AARG7818 standard; peptide; 14 AA.

XX

AC AARG7818;

XX

DT 18-AUG-1995 (first entry)

DE

DE Flk2 ligand N-terminal sequence.

XX

KW human Flk2 receptor protein-tyrosine-kinase ligand; peripheral

KW peripheral blood leukocyte conditioned medium;

KW bone marrow disorder diagnosis; hematopoietic stem cell;

KW proliferation; differentiation.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT misc_difference 7 /note= "any amino acid"

FT

XX WO9500554-A.

XX

PN 05-JAN-1995.

PD

XX

PF 17-JUN-1994; 94WO-US06944.

XX

XX 18-JUN-1993; 93US-0080244.

PR

XX 21-JUN-1993; 93US-0081508.

PR

XX 23-NOV-1993; 93US-0157490.

XX

PA (UYPR-) UNIV PRINCETON.

XX

XX Lemischka IR;

XX

XX WPI; 1995-052014/07.

XX

XX Ligand for receptor protein tyrosine kinase - useful for the

PT stimulation of primitive haematopoietic stem cells causing

PT proliferation and/or differentiation

XX

XX Claim 1; Page 105; 131pp; English.

XX

XX The sequence corresponds to the N-terminal region of a human Flk2

CC receptor protein-tyrosine-kinase ligand, isolated from

CC phytohemagglutinin-stimulated human peripheral blood leukocyte

CC tissue culture conditioned medium. The ligand may be used in

CC diagnosis of bone marrow disorders, and to stimulate the

CC proliferation and/or differentiation of primitive hematopoietic stem

CC cells. The ligand binds to a receptor protein-tyrosine-kinase

CC expressed in primitive but not mature mammalian hematopoietic cells.

XX

SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 16; Length 14;

Best Local Similarity 100.0%; Pred. No. 1.4e-05;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

DB 1 AQSLSFXFTKFDLD 14

RESULT 3

AARG2895

ID AARG2895 standard; peptide; 14 AA.

XX

AC AARG2895;

XX

DT 17-SEP-2001 (first entry)

DE

DE Antigenic peptide derived from a french bean FRIL polypeptide.

XX

KW FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;

KW progenitor cell preservation factor; radiotherapy; chemotherapy;

KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;

KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX

OS Phaseolus vulgaris.

XX

PH Key Location/Qualifiers

FT Misc-difference 7 /label= Asn, Cys, Ser

FT

XX WO200149851-A1.

XX

PD 12-JUL-2001.

XX

PF 30-DEC-1999; 99WO-US31307.

XX

PR 30-DEC-1999; 99WO-US31307.

XX

PA (PHYL-) PHYLOGIX LLC.

XX

XX Colucci MG, Chrispeels MJ, Moore JG;

XX WPI; 2001-441882/47.

XX

XX Legume Progenitor cell preservation factors for in vivo or ex vivo

PT preservation of hematopoietic progenitor cells and as therapeutics for

PT alleviating/reducing progenitor cell-depleting activity of cancer

XX therapeutics

XX

XX Example 5; Page 72; 173pp; English.

XX

XX The present sequence is derived from FRIL (Flk2/Flt3 tyrosine kinase

CC receptor-interacting lectin), and is used to raise antibodies. The

CC specification describes a composition of one or more members

CC of FRIL family of progenitor cell preservation factors. The composition

CC is useful for alleviating or reducing the hematopoietic progenitor

CC cell-depleting activity of a therapeutic treatment, including

CC radiotherapeutic and/or chemotherapeutic treatments. Administration of

CC FRIL compositions to a patient prior to treatment of the patient with

CC a therapeutic treatment having a hematopoietic progenitor cell-depleting

CC activity alleviates or reduces the hematopoietic progenitor

CC cell-depleting activity of the therapeutic treatment in the patient.

CC FRIL family members are useful for isolating population of progenitor

CC cells, hemangioblasts, and mesenchymal stem cells. The composition is

CC administered to reduce progenitor cell depleting effects of

chemotherapeutics, so that the patient can receive a higher dose of the chemotherapeutic and preferably recover from cancer. It is also administered to patients having, or predisposed to developing a condition where the patients hematopoietic progenitor cells are depleted, such as severe combined immunodeficiency or aplastic anemia. The isolated mesenchymal cells are useful for tissue repair.

XX SQ Sequence 14 AA;

Query Match 97.0%; Score 64; DB 22; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14
|||||

Db 1 AQSLSFXFTKFDLD 14

RESULT 4

AAW61497
ID AAW61497 standard; peptide; 12 AA.

XX AC AAW61497;

DT 19-OCT-1998 (first entry)

DE Pylartin protein, peptide chain beta (i).

XX Pylartin protein; progenitor cell; hematopoietic system; cancer;
KW engraftation; hematologic disease; sickle cell anemia; thalassemia.

XX OS Leguminosae.

XX PN W09825457-A1.

XX PD 18-JUN-1998.

XX PF 09-DEC-1997; 97WO-US22486.

XX PR 28-MAR-1997; 97US-0825369.

XX PS 09-DEC-1996; 96US-0762537.

XX PA (IMCL-) IMCLONE SYSTEMS INC.

XX PI Moore JG;

XX DR WPI; 1998-348161/30.

XX Pylartin proteins which preserve progenitor cells - useful for
PT hematopoietic therapies in cancer treatment or for treating
PT hematologic diseases

PS Claim 1; Page 28; 46pp; English.

XX The peptides AAW61497-W61502 can be used to form pylartin proteins which
CC preserve progenitor cells (pc). The proteins can be used for protecting
CC the integrity of the hematopoietic processes in vivo and as adjuncts in
CC therapeutic treatments related to cancer and other diseases which can
CC otherwise adversely impact upon the hematopoietic system. Since the
CC proteins bind specifically to primitive PCs, they can also be used for
CC the identification and localization of PCs. The methods can be used for
CC e.g. expanding PC populations ex vivo to increase chances of
CC engraftation, improving conditions for transporting and storing PCs and
CC for removing a fundamental barrier thereby enabling gene therapy to
CC treat and cure a broad range of life-threatening hematologic diseases
CC such as sickle cell anemia and thalassemia.

XX SQ Sequence 12 AA;

Query Match 83.3%; Score 55; DB 19; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00061;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFD 12
|||||

Db 1 AQSLSFXFTKFD 12

RESULT 5

AAG62896
ID AAG62896 standard; peptide; 14 AA.

XX AC AAG62896;

DT 17-SEP-2001 (first entry)

DE Peptide derived from a french bean FRIL polypeptide.

XX FRIL; Flk2/Flt3 tyrosine kinase receptor-interacting lectin;
KW progenitor cell preservation factor; radiotherapy; chemotherapy;
KW progenitor cell; hemangioblast; mesenchymal stem cell; cancer;
KW severe combined immunodeficiency; aplastic anemia; tissue repair.

XX OS Phaseolus vulgaris.

XX FH Key Location/Qualifiers

FT Misc-difference 7

FT /note= "not known"

PN W0200149851-A1.

XX PD 12-JUL-2001.

XX PF 30-DEC-1999; 99WO-US31307.

XX PR 30-DEC-1999; 99WO-US31307.

XX PA (PHYL-) PHYLOGIX LLC.

XX PI Colucci MG, Chrispeels MJ, Moore JG;

XX DR WPI; 2001-441882/47.

XX Legume progenitor cell preservation factors for in vivo or ex vivo
PT preservation of hematopoietic progenitor cells and as therapeutics for
PT alleviating/reducing progenitor cell-depleting activity of cancer
PT therapeutics

PS Example 5; Page 75; 173pp; English.

XX The present sequence is derived from a FRIL (Flk2/Flt3 tyrosine kinase
CC receptor-interacting lectin) polypeptide. The specification describes a
CC composition of one or more members of FRIL family of progenitor cell
CC preservation factors. The composition is useful for alleviating or
CC reducing the hematopoietic progenitor cell-depleting activity of a
CC therapeutic treatment, including radiotherapeutic and/or
CC chemotherapeutic treatments. Administration of FRIL compositions to a
CC patient prior to treatment of the patient with a therapeutic treatment
CC having a hematopoietic progenitor cell-depleting activity alleviates or
CC reduces the hematopoietic progenitor cell-depleting activity of the
CC therapeutic treatment in the patient. FRIL family members are useful for
CC isolating population of progenitor cells, hemangioblasts, and mesenchymal
CC stem cells. The composition is administered to reduce progenitor cell
CC depleting effects of chemotherapeutics, so that the patient can receive
CC a higher dose of the chemotherapeutic and preferably recover from cancer.
CC It is also administered to patients having, or predisposed to developing
CC a condition where the patients hematopoietic progenitor cells are
CC depleted, such as severe combined immunodeficiency or aplastic anemia.
CC The isolated mesenchymal cells are useful for tissue repair.

XX SQ Sequence 14 AA;

Query Match 71.2%; Score 47; DB 22; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.024;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
 DB 1 AQSLSFXFTKDALD 14

RESULT 6
 AAY62906
 ID AAY62906 standard; Peptide; 8 AA.
 XX
 AC AAY62906;
 XX
 DT 02-MAR-2000 (first entry)
 XX
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2462.
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..8 "the terminal residues are condensed with each
 FT /note= other to form a cyclic peptide"
 FT
 FT
 PN WO9957149-A2.
 XX
 PD 11-NOV-1999.
 XX
 XX 05-MAY-1999; 99WO-CA00363.
 XX
 PR 05-MAY-1998; 98US-0073040.
 PR 06-NOV-1998; 98US-0187859.
 PR 20-JAN-1999; 99US-0234395.
 PR 08-MAR-1999; 99US-0264516.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Blaschuk OW, Gour BJ, Byers S;
 XX
 DR WPI; 2000-038791/03.
 XX
 PT New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease -
 XX
 PS Claim 72; Page 195; 252pp; English.
 XX
 CC The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue

CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAY33183 to AAY33186 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 8 AA;
 Query Match 37.9%; Score 25; DB 21; Length 8;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KFDLD 14
 DB 1 KFDID 5
 |||:|
 RESULT 7
 AAY63026
 ID AAY63026 standard; Peptide; 8 AA.
 XX
 AC AAY63026;
 XX
 DT 02-MAR-2000 (first entry)
 XX
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2582.
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..8
 FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"
 FT
 FT
 PN WO9957149-A2.
 XX
 PD 11-NOV-1999.
 XX
 XX 05-MAY-1999; 99WO-CA00363.
 XX
 PR 05-MAY-1998; 98US-0073040.
 PR 06-NOV-1998; 98US-0187859.
 PR 20-JAN-1999; 99US-0234395.
 PR 08-MAR-1999; 99US-0264516.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Blaschuk OW, Gour BJ, Byers S;
 XX
 DR WPI; 2000-038791/03.
 XX
 PT New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease -
 XX
 PS Claim 72; Page 197; 252pp; English.
 XX
 CC The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue

CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 8 AA;

Query Match 37.9%; Score 25; DB 21; Length 8;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
 ||||
 Db 1 KFDID 5

RESULT 8

AAAY46829
 ID AAY46829 standard; Peptide; 9 AA.

XX AC AAY46829;

XX DT 01-DEC-1999 (first entry)

XX DE Immunogenic peptide having a human leukocyte antigen binding motif #1440.

XX KW Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
 KW immune response; T cell activation; major histocompatibility complex;
 KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
 KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
 KW vaccine; immunisation.

XX OS Synthetic.

OS Homo sapiens.

PN WO9945954-A1.

XX PD 16-SEP-1999.

XX PF 13-MAR-1998; 98WO-US05039.

XX PR 13-MAR-1998; 98WO-US05039.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;

XX DR WPI; 1999-551214/46.

XX PT New immunogenic peptides with HLA binding motif, useful in treatment
 PT and diagnosis of cancers and viral diseases

XX PS Claim 1; Page 86; 150pp; English.

XX CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also
 CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic
 CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes

CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell
 CC response against the antigen from which the peptide is derived.
 CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are
 CC normally induced by an antigen in the form of a peptide fragment bound
 CC to a HLA molecule, rather than the intact foreign antigen itself, and
 CC are particularly important in tumour rejection and in fighting viral
 CC infections. The peptides are therefore useful therapeutically to treat
 CC or prevent viral infections and cancers in mammals (especially humans)
 CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.
 CC They can be administered as vaccines to elicit an immune response in
 CC individuals susceptible or otherwise at risk of viral infection or
 CC cancer, or used to treat chronic or acute conditions. They are also
 CC useful diagnostically, and can be used to induce a cytotoxic T cell
 CC response, by contacting a cytotoxic T cell with the peptide e.g. to
 CC produce CTLs ex vivo for infusion back into a patient. The
 CC polynucleotides encoding the immunogenic peptides are also useful
 CC therapeutically and for immunisation as above.

XX Sequence 9 AA;

Query Match 37.9%; Score 25; DB 20; Length 9;
 Best Local Similarity 66.7%; Pred. No. 7.8e+05;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 QLSFXFTK 10
 |||||
 Db 1 QSSSFIFHK 9

RESULT 9

AAAY62907
 ID AAY62907 standard; Peptide; 9 AA.

XX AC AAY62907;

XX DT 02-MAR-2000 (first entry)

XX DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2463.

XX KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.

XX OS Synthetic.

OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT Modified-site 1..9

XX FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"

XX PN WO9957149-A2.

XX PD 11-NOV-1999.

XX PF 05-MAY-1999; 99WO-CA00363.

XX PR 05-MAY-1998; 98US-0073040.

XX PR 06-NOV-1998; 98US-0187859.

XX PR 20-JAN-1999; 99US-0234395.

XX PR 08-MAR-1999; 99US-0264516.

XX PA (ADHE-) ADHEREX TECHNOLOGIES INC.

XX PI Blaschuk OW, Gour BJ, Byers S;

XX DR WPI; 2000-038791/03.

PT New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease -
 XX
 PS Claim 72; Page 195; 252pp; English.
 XX
 CC The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 9 AA;
 Query Match 37.9%; Score 25; DB 21; Length 9;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KFDLD 14
 Db 1 KFDID 5
 RESULT 10
 AAY63027
 ID AAY63027 standard; Peptide; 9 AA.
 AC AAY63027;
 XX
 DT 02-MAR-2000 (first entry)
 XX
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2583.
 KW Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..9
 FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"
 XX
 PN W09957149-A2.
 XX
 PD 11-NOV-1999.
 XX

PF 05-MAY-1999; 99WO-CA00363.
 XX
 PR 05-MAY-1998; 98US-0073040.
 PR 06-NOV-1998; 98US-0187859.
 PR 20-JAN-1999; 99US-0234395.
 PR 08-MAR-1999; 99US-0264516.
 XX
 PA (ADHE-) ADHEREX TECHNOLOGIES INC.
 XX
 PI Blaschuk OW, Gour BJ, Byers S;
 XX
 DR WPI: 2000-038791/03.
 XX
 PT New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease -
 XX
 PS Claim 72; Page 197; 252pp; English.
 XX
 CC The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis, and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 9 AA;
 Query Match 37.9%; Score 25; DB 21; Length 9;
 Best Local Similarity 80.0%; Pred. No. 7.8e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KFDLD 14
 Db 1 KFDID 5
 RESULT 11
 AAB98514
 ID AAB98514 standard; peptide; 9 AA.
 XX
 AC AAB98514;
 XX
 DT 03-AUG-2001 (first entry)
 XX
 DE Human TAGD-15 peptide fragment #6.
 XX
 KW Human; TAGD-15; cytostatic; vaccine; ovarian tumour; cancer;
 KW tumour antigen-derived gene 15; serine protease.
 XX
 OS Homo sapiens.
 OS W0200129056-A1.
 XX
 PD 26-APR-2001.
 XX

XX PF 20-OCT-2000; 2000WO-US29095.
 XX PR 20-OCT-1999; 99US-0421213.
 XX PA (UYAR-) UNIV ARKANSAS.
 XX PI O'Brien TJ, Tanimoto H;
 XX DR WPI; 2001-381031/40.
 XX PT Novel extracellular serine protease, termed tumor antigen-derived gene
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for
 PT diagnosis, treatment, prevention of cancer, particularly breast,
 PT ovarian cancer -
 XX PS Example 9; Page 43; 130pp; English.
 XX CC The present invention relates to human tumour antigen-derived gene 15
 CC (TAGD-15) protein and coding sequence (see AAH23601 and AAB98500).
 CC TAGD-15 is an extracellular serine protease. It was found that TAGD-15 is
 CC over-expressed in ovarian tumours. TAGD-15 protein or its fragments of
 CC 9-20 residues that lack TAGD-15 protease activity are useful for
 CC vaccinating an individual against TAGD-15, having, suspected of having or
 CC at risk of getting cancer. The present sequence is one such peptide
 CC fragment of TAGD-15.
 XX SQ Sequence 9 AA;
 Query Match 37.9%; Score 25; DB 22; Length 9;
 Best Local Similarity 55.6%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 5 SFXTKFDL 13
 Db 1 SLTFRSFDL 9
 RESULT 12
 AAB98560
 ID AAB98560 standard; peptide; 9 AA.
 AC AAB98560;
 DT 03-AUG-2001 (first entry)
 XX Human TAGD-15 peptide fragment #52.
 DE Human; TAGD-15; cytostatic; vaccine; ovarian tumour; cancer;
 KW tumour antigen-derived gene 15; serine protease.
 XX Homo sapiens.
 OS WO200129056-A1.
 PN 26-APR-2001.
 XX 20-OCT-2000; 2000WO-US29095.
 PF 20-OCT-1999; 99US-0421213.
 XX (UYAR-) UNIV ARKANSAS.
 PA O'Brien TJ, Tanimoto H;
 XX WPI; 2001-381031/40.
 XX PT Novel extracellular serine protease, termed tumor antigen-derived gene
 PT 15 protein overexpressed in carcinomas and DNA encoding it, for
 PT diagnosis, treatment, prevention of cancer, particularly breast,
 PT ovarian cancer -
 XX PS Example 9; Page 44; 130pp; English.

XX CC The present invention relates to human tumour antigen-derived gene 15
 CC (TAGD-15) protein and coding sequence (see AAH23601 and AAB98500).
 CC TAGD-15 is an extracellular serine protease. It was found that TAGD-15 is
 CC over-expressed in ovarian tumours. TAGD-15 protein or its fragments of
 CC 9-20 residues that lack TAGD-15 protease activity are useful for
 CC vaccinating an individual against TAGD-15, having, suspected of having or
 CC at risk of getting cancer. The present sequence is one such peptide
 CC fragment of TAGD-15.
 XX SQ Sequence 9 AA;
 Query Match 37.9%; Score 25; DB 22; Length 9;
 Best Local Similarity 55.6%; Pred. No. 7.8e+05;
 Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 5 SFXTKFDL 13
 Db 1 SLTFRSFDL 9
 RESULT 13
 AAY78322
 ID AAY78322 standard; Protein; 10 AA.
 AC AAY78322;
 XX 04-MAY-2000 (first entry)
 DT Anti-zeta-chain antibody 2-B-5 VH-region CDR1 protein sequence.
 DE Anti-zeta-chain antibody 2-B-5; rat; human zeta chain; immunisation;
 KW complementary determining region; CDR; autoimmune disease; cytostatic;
 KW immune deficiency; T-cell malignancy; infectious disease; antiviral;
 KW immunosuppressive; antimicrobial; immune response modulator; NK-cell.
 XX Rattus norvegicus.
 OS WO200003016-A1.
 PN 20-JAN-2000.
 XX 09-JUL-1999; 99WO-EP04838.
 PF 10-JUL-1998; 98EP-0112867.
 PR (CONN-) CONNEX GMBH.
 XX Reiter C;
 PI WPI; 2000-160926/14.
 DR N-PSDB; AAZ88320.
 XX New oligonucleotide, polypeptide, antibody useful for treating
 PT autoimmune disease, immune deficiencies, T-cell malignancies and
 PT infectious diseases -
 XX Claim 10; Fig 6; 79pp; English.
 XX The present invention describes a nucleic acid molecule (I) encoding at
 CC least one complementary determining region (CDR) of a variable region of
 CC an antibody which specifically interacts with the extracellular domain of
 CC the human zeta-chain. The antibody whose CDR of a variable region is
 CC encoded by (I), is obtained by immunising a rat with Jurkat cells and
 CC subsequently with a conjugate comprising a carrier molecule and a
 CC peptide having 11 N-terminal amino acids of a rat zeta-chain. The
 CC anti-zeta-chain antibody is useful for the treatment and prevention of
 CC autoimmune diseases, immune deficiencies, T-cell malignancies,
 CC infectious diseases and the suppression of immune response preferably in
 CC order to avoid graft rejection after organ transplantation, malignancies,
 CC or viral infections. The antibody, and fragments of it, can be useful for
 CC the enhancement or suppression of NK-cell dependent immunity or for the
 CC treatment of NK-cell derived malignancies. It can also be useful for the

CC determination of zeta-chain or eta-chain expression on NK-cells,
 CC T-lymphocytes or their precursors. The present sequence represents the
 CC CDR1 of the VH-region of the anti-zeta-chain antibody 2-B-5, produced
 CC by rats from the present invention.

XX Sequence 10 AA;
 Query Match 37.9%; Score 25; DB 21; Length 10;
 Best Local Similarity 37.5%; Pred. No. 2.4e+02;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXFTKFDL 13
 Db 2 YFTSYDM 9

RESULT 14
 AAY62908
 ID AAY62908 standard; Peptide; 10 AA.

XX AAY62908;
 DT 02-MAR-2000 (first entry)
 DE PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2464.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.

XX Synthetic.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 1.10
 FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"

XX WO957149-A2.
 XX 11-NOV-1999.
 XX 05-MAY-1999; 99WO-CA00363.
 XX 05-MAY-1998; 98US-0073040.
 XX 06-NOV-1998; 98US-0187859.
 XX 20-JAN-1999; 99US-0234395.
 XX 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

XX New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease

XX Claim 72; Page 195; 252pp; English.

XX The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MA can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,

CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 10 AA;

Query Match 37.9%; Score 25; DB 21; Length 10;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
 Db 1 KFDID 5

RESULT 15

AAY63028
 ID AAY63028 standard; Peptide; 10 AA.

XX AAY63028;

XX 02-MAR-2000 (first entry)

XX PB-cadherin cell adhesion recognition cyclic peptide SEQ ID NO:2584.

XX Modulation; nonclassical cadherin mediated cell adhesion; CAR;
 KW inhibition; cadherin extracellular domain; cell adhesion recognition;
 KW OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8;
 KW cadherin-12; cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
 KW cadherin related neuronal receptor; LI-cadherin; protocadherin;
 KW desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
 KW rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
 KW neurological disease; cyclic.

XX Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 1.10
 FT /note= "the terminal residues are condensed with each
 FT other to form a cyclic peptide"

XX WO957149-A2.

XX 11-NOV-1999.

XX 05-MAY-1999; 99WO-CA00363.

XX 05-MAY-1998; 98US-0073040.

XX 06-NOV-1998; 98US-0187859.

XX 20-JAN-1999; 99US-0234395.

XX 08-MAR-1999; 99US-0264516.

XX (ADHE-) ADHEREX TECHNOLOGIES INC.

XX Blaschuk OW, Gour BJ, Byers S;

XX WPI; 2000-038791/03.

PT New cadherin modulating agents, used for modulating nonclassical
 PT cadherin-mediated functions for treating e.g. cancers, obesity,
 PT rheumatoid arthritis, multiple sclerosis, diabetes or a neurological
 PT disease

XX
 PS Claim 72; Page 197; 252pp; English.
 XX

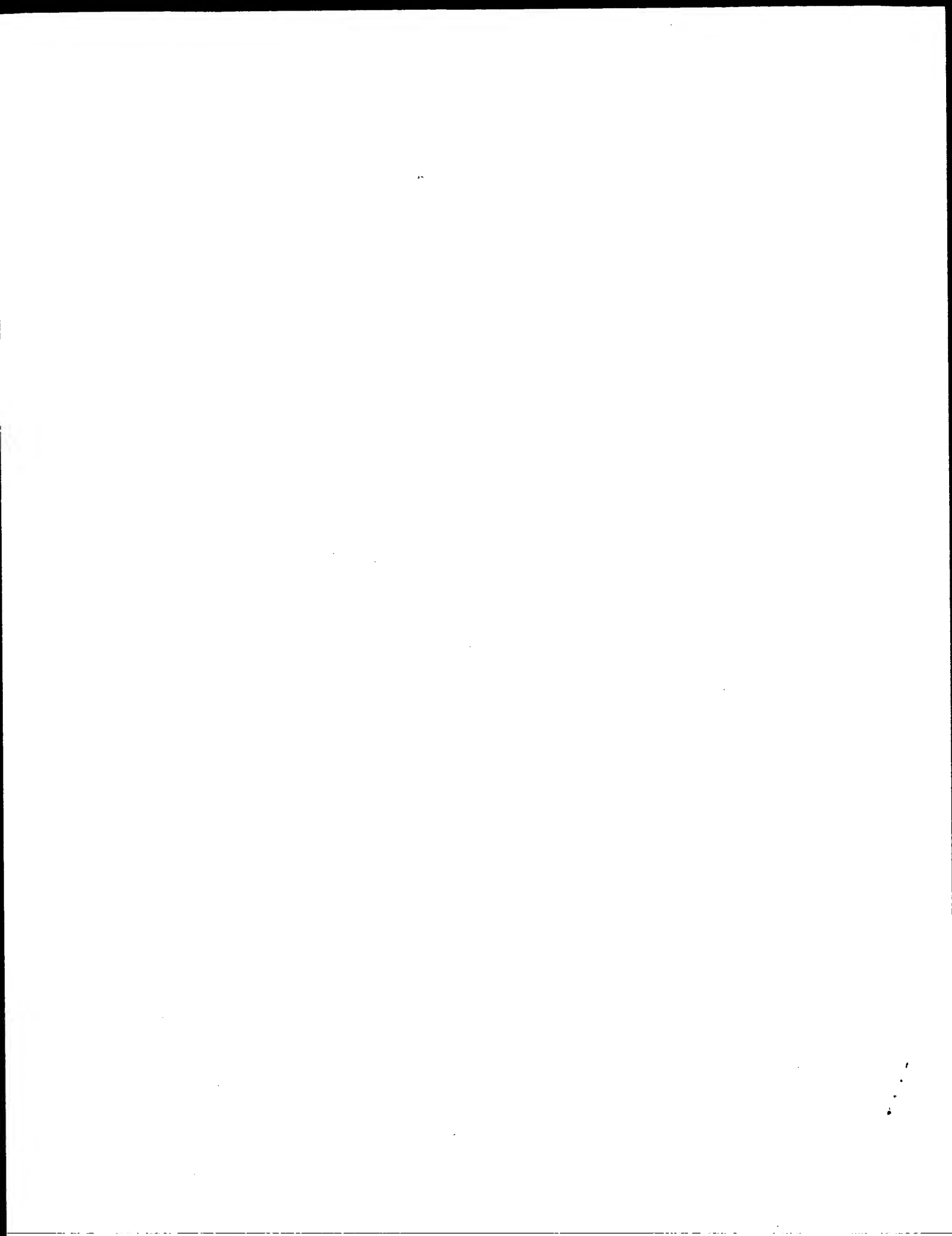
CC The present invention describes cadherin modulating agents (MA)
 CC comprising peptides which comprise a nonclassical cadherin cell adhesion
 CC recognition (CAR) sequence. The MAs can be used for modulating
 CC nonclassical cadherin-mediated functions. They can be used for e.g.
 CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
 CC mammal, enhancing delivery of a drug through the skin of a mammal,
 CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
 CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
 CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
 CC expressing cell, preventing or treating obesity in a mammal, stimulating
 CC blood vessel regression in a mammal, enhancing drug delivery to the
 CC central nervous system, treating a demyelinating neurological disease,
 CC increasing vasopermeability in a mammal, enhancing adhesion of
 CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
 CC a mammal, or preventing pregnancy in a mammal. They can also be used for
 CC e.g. enhancing or directing neurite outgrowth, facilitating wound
 CC healing or reducing scar tissue, or enhancing adhesion of foreign tissue
 CC in a mammal. They can also be used for treating e.g. psoriasis,
 CC arthritis, age-related macular degeneration, multiple sclerosis and
 CC diabetes. The products can also be used for detection and diagnosis and
 CC in bioreactors. AAY60592 to AAY64572 represent specifically claimed
 CC peptides, and AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent
 CC sequences used in the exemplification of the present invention.
 XX

SQ Sequence 10 AA;

Query Match 37.9%; Score 25; DB 21; Length 10;
 Best Local Similarity 80.0%; Pred. NO. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
 |||:
 Db 1 KFDID 5

Search completed: February 26, 2003, 14:56:22
 Job time : 35 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:57:21 : Search time 12 Seconds
(without alignments)
44.009 Million cell updates/sec

Title: US-09-476-485A-31
Perfect score: 66
Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174566 seqs, 37721826 residues

Total number of hits satisfying chosen parameters: 32941

Minimum DB seq length: 0
Maximum DB seq length: 14

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PublishedApplications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	98.5	14	10	US-09-934-251A-2
2	55	83.3	12	10	US-09-934-251A-1
3	25	37.9	13	9	US-09-896-896A-65
4	25	37.9	13	10	US-09-753-126-101
5	24	36.4	9	10	US-09-812-079A-7
6	24	36.4	10	9	US-09-900-590-34
7	24	36.4	10	9	US-09-791-389-249
8	24	36.4	10	9	US-09-791-393-249
9	23	34.8	9	10	US-09-780-053-43
10	23	34.8	9	10	US-09-780-053-230
11	23	34.8	9	10	US-09-780-053-317
12	23	34.8	10	10	US-09-780-053-83
13	23	34.8	10	10	US-09-780-053-286
14	23	34.8	10	10	US-09-780-053-391
15	23	34.8	13	9	US-10-075-846-18
16	23	34.8	13	10	US-09-870-379-12
17	22	33.3	14	9	US-09-974-879-475
18	22	33.3	10	9	US-09-900-590-52
19	22	33.3	11	9	US-09-921-650-4

20	22	33.3	11	10	US-09-874-389-4	Sequence 4, Appli
21	22	33.3	12	10	US-09-832-723-18	Sequence 18, Appl
22	22	33.3	13	9	US-09-826-290-359	Sequence 359, App
23	22	33.3	13	9	US-09-826-290-372	Sequence 372, App
24	22	33.3	13	9	US-09-791-389-116	Sequence 116, App
25	22	33.3	13	9	US-09-791-393-116	Sequence 116, App
26	22	33.3	13	10	US-09-791-378-213	Sequence 213, App
27	22	33.3	13	10	US-09-791-378-223	Sequence 223, App
28	21	31.8	9	9	US-10-032-482-14	Sequence 14, Appl
29	21	31.8	10	10	US-09-780-053-278	Sequence 278, App
30	21	31.8	10	10	US-09-780-053-709	Sequence 709, App
31	21	31.8	14	9	US-09-764-868-1242	Sequence 1242, Ap
32	21	31.8	14	9	US-09-955-999-122	Sequence 122, App
33	20	30.3	8	9	US-10-079-625-33	Sequence 33, Appl
34	20	30.3	9	8	US-08-737-457A-8	Sequence 8, Appli
35	20	30.3	9	9	US-10-106-487-21	Sequence 21, Appl
36	20	30.3	9	9	US-09-277-074-8	Sequence 8, Appli
37	20	30.3	9	9	US-09-791-389-166	Sequence 166, App
38	20	30.3	9	9	US-09-791-393-166	Sequence 166, App
39	20	30.3	9	10	US-09-759-960-24	Sequence 24, Appl
40	20	30.3	9	10	US-09-789-720-2	Sequence 2, Appli
41	20	30.3	9	10	US-09-124-280A-40	Sequence 40, Appl
42	20	30.3	9	10	US-09-950-313-23	Sequence 23, Appl
43	20	30.3	9	10	US-09-780-053-619	Sequence 619, App
44	20	30.3	9	10	US-09-912-787-84	Sequence 84, Appl
45	20	30.3	9	10	US-09-756-983-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-934-251A-2
; Sequence 2, Application US/09934251A
; Patent No. US20020132017A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934, 251A
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

Query Match 98.5%; Score 65; DB 10; Length 14;
Best Local Similarity 92.9%; Pred. No. 3.2e-06;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
| | | | | | | | | | | | | | | |
Db 1 AQSLSFXFTKFDLD 14

RESULT 2

US-09-934-251A-1
; Sequence 1, Application US/09934251A
; Patent No. US20020132017A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT APPLICATION NUMBER: US/09/934, 251A

```
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide chain of pylartin protein
US-09-934-251A-1

Query Match      83.3%; Score 55; DB 10; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0002;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12
   ||||| |||||
Db 1 AQSLSFXTKFD 12

RESULT 3
US-09-896-896A-65
; Sequence 65, Application US/09896896A
; Publication No. US20030036181A1
; GENERAL INFORMATION:
; APPLICANT: MAXYGEN APS
; TITLE OF INVENTION: PEPTIDE EXTENDED GLYCOSYLATED POLYPEPTIDES
; FILE REFERENCE: 0217us210
; CURRENT APPLICATION NUMBER: US/09/896,896A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/217,497
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: US 60/225,558
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: DK PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DK PA 2000 01092
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: PCT/DK00/00743
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: PCT/DK01/00090
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 65
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-896-896A-65

Query Match      37.9%; Score 25; DB 9; Length 13;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXT 9
   ||: |||
Db 1 AQNKTENFT 9

RESULT 4
US-09-753-126-101
; Sequence 101, Application US/09753126
; Patent No. US20020127219A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
```

```
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-000600US
; CURRENT APPLICATION NUMBER: US/09/753,126
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-753-126-101

Query Match      37.9%; Score 25; DB 10; Length 13;
Best Local Similarity 55.6%; Pred. No. 80;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXT 9
   ||: |||
Db 1 AQNKTENFT 9

RESULT 5
US-09-812-079A-7
; Sequence 7, Application US/09812079A
; Patent No. US20020058038A1
; GENERAL INFORMATION:
; APPLICANT: Nicolette, Charles A.
; TITLE OF INVENTION: THERAPEUTIC ANTI-CYTOMEGALOVIRUS
; TITLE OF INVENTION: COMPOUNDS
; FILE REFERENCE: GZ 209500
; CURRENT APPLICATION NUMBER: US/09/812,079A
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: 60/191,050
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/254,989
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-812-079A-7

Query Match      36.4%; Score 24; DB 10; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.5e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 FXTKTF 11
   |||:|
Db 1 FLFTRF 6
```


RESULT 6

US-09-900-590-34
; Sequence 34, Application US/09900590
; Publication No. US20030028009A1
; GENERAL INFORMATION:
; APPLICANT: Huse, William D.
; Glaser, Scott M.
; TITLE OF INVENTION: Anti-Alpha V Beta 3 Recombinant Human
; Antibodies, Nucleic Acids Encoding Same and Methods of Use
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/900,590
; FILING DATE: 06-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/016,061
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IX 2965
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-900-590-34

Query Match 36.4%; Score 24; DB 9; Length 10;
Best Local Similarity 37.5%; Pred. No. 94;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXFTKFDL 13
| | : | :
Db 2 FTFSYDM 9

RESULT 7

US-09-791-389-249
; Sequence 249, Application US/09791389
; Publication No. US20030032773A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Rohlf, Christian
; APPLICANT: Terrett, Jonathan Alexander
; APPLICANT: Tyson, Kerry Louise
; TITLE OF INVENTION: Proteins, Genes and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
; TITLE OF INVENTION: and Unipolar Depression
; FILE REFERENCE: 2543-1-001 N2
; CURRENT APPLICATION NUMBER: US/09/791,389
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: GB 0004412.3
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: GB 0030050.9

; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/254,830
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapien
US-09-791-389-249

Query Match 36.4%; Score 24; DB 9; Length 10;
Best Local Similarity 62.5%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12
| | | | |
Db 1 SADFTNFD 8

RESULT 8

US-09-791-393-249
; Sequence 249, Application US/09791393
; Publication No. US20030032200A1
; GENERAL INFORMATION:
; APPLICANT: Herath, Herath Mudiyanseelage Athula Chandrasiri
; APPLICANT: Parekh, Rajesh Bhikhu
; APPLICANT: Rohlf, Christian
; TITLE OF INVENTION: Proteins, Genes and Their Use for
; TITLE OF INVENTION: Diagnosis and Treatment of Bipolar Affective Disorder (BAD)
; TITLE OF INVENTION: and Unipolar Depression
; FILE REFERENCE: 2543-1-001 N1
; CURRENT APPLICATION NUMBER: US/09/791,393
; CURRENT FILING DATE: 2002-01-02
; EARLIER APPLICATION NUMBER: GB 0004412.3
; EARLIER FILING DATE: 2000-02-24
; EARLIER APPLICATION NUMBER: GB 0030050.9
; EARLIER FILING DATE: 2000-12-08
; EARLIER APPLICATION NUMBER: US 60/254,830
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 249
; LENGTH: 10
; TYPE: PRT
; ORGANISM: homo sapien
US-09-791-393-249

Query Match 36.4%; Score 24; DB 9; Length 10;
Best Local Similarity 62.5%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12
| | | | |
Db 1 SADFTNFD 8

RESULT 9

US-09-780-053-43
; Sequence 43, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Chailita-Eld
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.50SUI
; CURRENT APPLICATION NUMBER: US/09/780,053

; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-43

Query Match 34.8%; Score 23; DB 10; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12
| |:
Db 1 SVAFSKFE 8

RESULT 10
US-09-780-053-230
; Sequence 230, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 230
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-230

Query Match 34.8%; Score 23; DB 10; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12
| |:
Db 1 SVAFSKFE 8

RESULT 11
US-09-780-053-317
; Sequence 317, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 317
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-317

Query Match 34.8%; Score 23; DB 10; Length 9;
Best Local Similarity 50.0%; Pred. No. 1.5e+05;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12
| |:
Db 1 SVAFSKFE 8

RESULT 12
US-09-780-053-83
; Sequence 83, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-83

Query Match 34.8%; Score 23; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 SFXFTKFD 12
| |:
Db 2 SVAFSKFE 9

RESULT 13
US-09-780-053-286
; Sequence 286, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261

; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-286

Query Match 34.8%; Score 23; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12
| | | | |
Db 2 SVAFSKFE 9

RESULT 14

US-09-780-053-391
; Sequence 391, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Ava Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5US01
; CURRENT APPLICATION NUMBER: US/09780,053
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-391

Query Match 34.8%; Score 23; DB 10; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 SFXFTKFD 12
| | | | |
Db 2 SVAFSKFE 9

RESULT 15

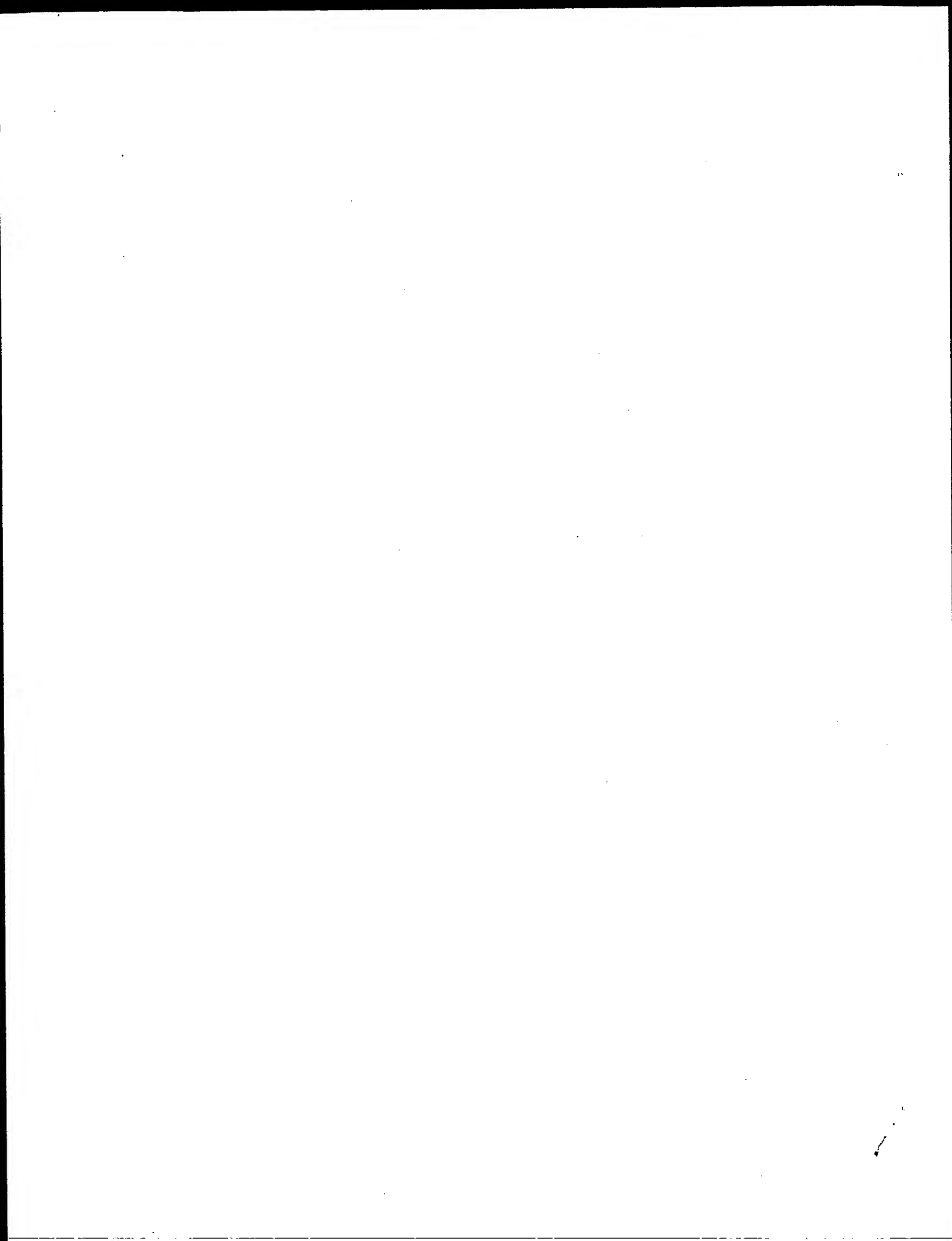
US-10-075-846-18
; Sequence 18, Application US/10075846
; Publication No. US20030032608A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING A NOVEL GLYCINE RECEPTOR ALPHA SUBUNIT
; TITLE OF INVENTION: IN THE GASTROINTESTINAL TRACT, HGRA4, and SPLICE VARIANT THEREOF
; FILE REFERENCE: D0079 NP
; CURRENT APPLICATION NUMBER: US/10/075,846
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US 60/269,535
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 18
; LENGTH: 13
; TYPE: PRT
; ORGANISM: homo sapiens

US-10-075-846-18

Query Match 34.8%; Score 23; DB 9; Length 13;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
| | | | |
Db 2 SFGYTMKDL 10

Search completed: February 26, 2003, 15:01:35
Job time : 13 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:56:26 ; Search time 132 seconds
(without alignments)

68.381 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 509362

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Parents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS06_COMB.pep.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
- 24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
- 25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
- 26: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
- 27: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	65	98.5	14	1	PCT-US97-22486-2
2	65	98.5	14	23	US-09-934-251A-1
3	64	97.0	14	4	US-08-081-508-12
4	64	97.0	14	5	US-08-157-490-11
5	64	97.0	14	18	US-09-476-485A-31
6	55	83.3	12	1	PCT-US97-22486-1

7	55	83.3	12	23	US-09-934-251A-1	Sequence 1, Appli
8	47	71.2	14	18	US-09-476-485A-32	Sequence 32, Appl
9	38.5	58.3	13	18	US-09-476-485A-34	Sequence 34, Appl
10	28	42.4	12	13	US-08-905-825-67	Sequence 67, Appl
11	27	40.9	6	4	US-08-080-244-12	Sequence 12, Appl
12	25	37.9	8	22	US-09-839-542-2462	Sequence 2462, Ap
13	25	37.9	8	22	US-09-839-542-2582	Sequence 2582, Ap
14	25	37.9	8	22	US-09-839-542B-2462	Sequence 2462, Ap
15	25	37.9	8	22	US-09-839-542B-2582	Sequence 2582, Ap
16	25	37.9	8	24	US-10-006-869-2462	Sequence 2462, Ap
17	25	37.9	8	24	US-10-006-869-2582	Sequence 2582, Ap
18	25	37.9	9	1	PCT-US00-29095-24	Sequence 24, Appl
19	25	37.9	9	1	PCT-US00-29095-70	Sequence 70, Appl
20	25	37.9	9	9	US-08-589-107-120	Sequence 120, App
21	25	37.9	9	18	US-09-421-213-24	Sequence 24, Appl
22	25	37.9	9	18	US-09-421-213-70	Sequence 70, Appl
23	25	37.9	9	20	US-09-654-600A-24	Sequence 24, Appl
24	25	37.9	9	20	US-09-654-600A-70	Sequence 70, Appl
25	25	37.9	9	22	US-09-839-542-2463	Sequence 2463, Ap
26	25	37.9	9	22	US-09-839-542-2583	Sequence 2583, Ap
27	25	37.9	9	22	US-09-839-542B-2463	Sequence 2463, Ap
28	25	37.9	9	22	US-09-839-542B-2583	Sequence 2583, Ap
29	25	37.9	9	24	US-10-006-869-2463	Sequence 2463, Ap
30	25	37.9	9	24	US-10-006-869-2583	Sequence 2583, Ap
31	25	37.9	10	21	US-09-743-482A-8	Sequence 8, Appli
32	25	37.9	10	22	US-09-839-542-2464	Sequence 2464, Ap
33	25	37.9	10	22	US-09-839-542-2584	Sequence 2584, Ap
34	25	37.9	10	22	US-09-839-542B-2464	Sequence 2464, Ap
35	25	37.9	10	22	US-09-839-542B-2584	Sequence 2584, Ap
36	25	37.9	10	24	US-10-006-869-2464	Sequence 2464, Ap
37	25	37.9	10	24	US-10-006-869-2584	Sequence 2584, Ap
38	25	37.9	11	23	US-09-977-827-10	Sequence 10, Appl
39	25	37.9	12	1	PCT-US00-07946-273	Sequence 273, App
40	25	37.9	12	19	US-09-534-717-273	Sequence 273, App
41	25	37.9	12	19	US-09-540-018-31	Sequence 31, Appl
42	25	37.9	12	22	US-09-801-185A-31	Sequence 31, Appl
43	25	37.9	12	22	US-09-801-185B-31	Sequence 31, Appl
44	25	37.9	12	25	US-10-133-715-31	Sequence 31, Appl
45	25	37.9	13	21	US-09-753-126-101	Sequence 101, App

ALIGNMENTS

RESULT 1
PCT-US97-22486-2
; Sequence 2, Application PC/TUS9722486
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/22486
; FILING DATE: 9-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/762,537
; FILING DATE: 9-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/825,369
FILING DATE: 28-MAR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Dea, Sean W.
REGISTRATION NUMBER: 37690
REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US97-22486-2

Query Match 98.5%; Score 65; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14
Db 1 AQSLSFXTKFDLD 14

RESULT 2

US-09-934-251A-2

Sequence 2, Application US/09934251A
GENERAL INFORMATION:
APPLICANT: Moore, Jeffrey G.
TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
FILE REFERENCE: 108236.136
CURRENT APPLICATION NUMBER: US/09/934,251A
CURRENT FILING DATE: 2001-08-21
PRIOR APPLICATION NUMBER: US 09/368,607
PRIOR FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: US 08/762,537
PRIOR FILING DATE: 1996-12-09
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 14
TYPE: PPT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: beta peptide sequence
US-09-934-251A-2

Query Match 98.5%; Score 65; DB 23; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.0001;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14
Db 1 AQSLSFXTKFDLD 14

RESULT 3

US-08-081-508-12

Sequence 12, Application US/08081508
GENERAL INFORMATION:
APPLICANT: Lemischka, Ihor R.
TITLE OF INVENTION: TORIPOINT HEMATOPOIETIC STEM CELL
TITLE OF INVENTION: RECEPTORS AND THEIR LIGANDS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: ImClone Systems Incorporated
STREET: 180 Varick Street
CITY: New York
STATE: New York
COUNTRY: U.S.A.

ZIP: 10014
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/081,508
FILING DATE: 19930621
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/679,666
FILING DATE: 02-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/728,913
FILING DATE: 28-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,065
FILING DATE: 15-NOV-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/813,593
FILING DATE: 24-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/906,397
FILING DATE: 26-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/975,049
FILING DATE: 12-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/977,451
FILING DATE: 19-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/005,941
FILING DATE: 15-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,272
FILING DATE: 01-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US UNASSIGNED
FILING DATE: 09-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US UNASSIGNED
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Felt, Irving N.
REGISTRATION NUMBER: 28,601
REFERENCE/DOCKET NUMBER: LEM-3-12P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-645-1405
TELEFAX: 212-645-2054
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-081-508-12

Query Match 97.0%; Score 64; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFDLD 14
Db 1 AQSLSFXTKFDLD 14

RESULT 4
US-08-157-490-11

; Sequence 11, Application US/08157490
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; TOPIC OF INVENTION: RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ImClone Systems Incorporated
; STREET: 180 Varick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,490
; FILING DATE: 23-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/975,049
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,941
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,272
; FILING DATE: 01-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/076022
; FILING DATE: 09-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/080244
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/081508
; FILING DATE: 21-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/096759
; FILING DATE: 22-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/125669
; FILING DATE: 23-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-15P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-157-490-11

Query Match 97.0%; Score 64; DB 5; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14
Db 1 AQSLSFXFTKFDLD 14
|||||

RESULT 5

US-09-476-485A-31
; Sequence 31, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; TITLE OF INVENTION: and Products of Their Use
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide corresponding to Pv-FRIL.
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = Asn, Cys or Ser.
; US-09-476-485A-31

Query Match 97.0%; Score 64; DB 18; Length 14;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AQSLSFXFTKFDLD 14
Db 1 AQSLSFXFTKFDLD 14
|||||

RESULT 6

PCT-US97-22486-1
; Sequence 1, Application PC/TUS9722486
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR
; TITLE OF INVENTION: PRESERVING PROGENITOR CELLS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage

; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/22486
; FILING DATE: 9-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA: 08/762,537
; FILING DATE: 9-DEC-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/825,369
; FILING DATE: 28-MAR-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 381-21 CIP/PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US97-22486-1

Query Match 83.3%; Score 55; DB 1; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
||||| |||||
Db 1 AQSLSFSFTKFD 12

RESULT 7
US-09-934-251A-1

; Sequence 1, Application US/09934251A
; GENERAL INFORMATION:
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Composition and Method for Preserving Progenitor Cells
; FILE REFERENCE: 108236.136
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 09/368,607
; PRIOR FILING DATE: 1999-08-05
; PRIOR APPLICATION NUMBER: US 08/762,537
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide chain of pylartin protein
US-09-934-251A-1

Query Match 83.3%; Score 55; DB 23; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.0063;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFD 12
||||| |||||
Db 1 AQSLSFSFTKFD 12

RESULT 8
US-09-476-485A-32

; Sequence 32, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; FILING DATE: 9-DEC-1997
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide of 18 kDa.
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = unknown amino acid.
US-09-476-485A-32

Query Match 71.2%; Score 47; DB 18; Length 14;
Best Local Similarity 85.7%; Pred. No. 0.22;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14
||||| ||||| ||
Db 1 AQSLSFXFTKDALD 14

RESULT 9

US-09-476-485A-34
; Sequence 34, Application US/09476485A
; GENERAL INFORMATION:
; APPLICANT: Colucci, M. Gabriella
; APPLICANT: Chrispeels, Maarten J.
; APPLICANT: Moore, Jeffrey G.
; TITLE OF INVENTION: Progenitor Cell Preservation Factors and Methods for
; FILING DATE: 9-DEC-1997
; FILE REFERENCE: 108236.119
; CURRENT APPLICATION NUMBER: US/09/476,485A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 08/881,189
; PRIOR FILING DATE: 1997-06-24
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Aminoterminal polypeptide.
; NAME/KEY: PEPTIDE
; LOCATION: (7)..(7)
; OTHER INFORMATION: Amino acid 7 is Xaa wherein Xaa = unknown amino acid.
US-09-476-485A-34

Query Match 58.3%; Score 38.5; DB 18; Length 13;
Best Local Similarity 91.7%; Pred. No. 7.7;
Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 AQSLSFXFTKFD 12
||||| ||||| ||
Db 1 AQSLSFXF-KFD 11

RESULT 10


```

US-08-905-825-67
; Sequence 67, Application US/08905825
; GENERAL INFORMATION:
; APPLICANT: Watkins, Jeffrey D.
; APPLICANT: Huse, William D.
; APPLICANT: Wu, Heiren
; TITLE OF INVENTION: METHODS FOR IDENTIFYING LIGAND SPECIFIC
; BINDING MOLECULES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,825
; FILING DATE: 04-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-IX 1570
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-905-825-67

```

```

Query Match 42.4%; Score 28; DB 13; Length 12;
Best Local Similarity 40.0%; Pred. No. 6.2e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 SPXTKFDLD 14
   |||: |||
Db 3 TTFESNYDLN 12

```

```

RESULT 11
US-08-080-244-12
; Sequence 12, Application US/08080244
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor R.
; TITLE OF INVENTION: TOTIPOTENT HEMATOPOIETIC STEM CELL
; RECEPTORS AND THEIR LIGANDS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Imclone Systems Incorporated
; STREET: 180 Varlick Street
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10014
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/080,244
; FILING DATE: 19930618
; CLASSIFICATION: 530

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/679,666
; FILING DATE: 02-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,913
; FILING DATE: 28-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/793,065
; FILING DATE: 15-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/813,593
; FILING DATE: 24-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/906,397
; FILING DATE: 26-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/975,049
; FILING DATE: 12-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,451
; FILING DATE: 19-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/005,941
; FILING DATE: 15-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,272
; FILING DATE: 01-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US UNASSIGNED
; FILING DATE: 09-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Feit, Irving N.
; REGISTRATION NUMBER: 28,601
; REFERENCE/DOCKET NUMBER: LEM-3-11P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-645-1405
; TELEFAX: 212-645-2054
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-080-244-12

```

```

Query Match 40.9%; Score 27; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AQSLSF 6
   |||||
Db 1 AQSLSF 6

```

```

RESULT 12
US-09-839-542-2462
; Sequence 2462, Application US/09839542
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2462

```

; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-839-542-2462

Query Match 37.9%; Score 25; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
|||:|
Db 1 KFDID 5

RESULT 13
US-09-839-542-2582
; Sequence 2582, Application US/09839542
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2582
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-839-542-2582

Query Match 37.9%; Score 25; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
|||:|
Db 1 KFDID 5

RESULT 14
US-09-839-542B-2462
; Sequence 2462, Application US/09839542B
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2462
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-839-542B-2462

Query Match 37.9%; Score 25; DB 22; Length 8;

Best Local Similarity 80.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
|||:|
Db 1 KFDID 5

RESULT 15
US-09-839-542B-2582
; Sequence 2582, Application US/09839542B
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Symonds, James Matthew
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; FILE REFERENCE: 100086.407D1
; CURRENT APPLICATION NUMBER: US/09/839,542B
; CURRENT FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2582
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-839-542B-2582

Query Match 37.9%; Score 25; DB 22; Length 8;
Best Local Similarity 80.0%; Pred. No. 4.2e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
|||:|
Db 1 KFDID 5

Search completed: February 26, 2003, 15:00:47
Job time : 133 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:56:46 ; Search time 23 seconds
(without alignments)
55.633 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 486122 seqs, 91396495 residues

Total number of hits satisfying chosen parameters: 152502

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	39.4	9	5	US-09-641-528-13672
2	26	39.4	9	5	US-09-641-528-45663
3	26	39.4	9	5	US-09-641-528A-13672
4	26	39.4	9	5	US-09-641-528A-45663
5	26	39.4	10	5	US-09-641-528-21738
6	26	39.4	10	5	US-09-641-528A-21738
7	26	39.4	11	5	US-09-641-528-18763
8	26	39.4	11	5	US-09-641-528A-18763
9	24	36.4	9	5	US-09-641-528-16119
10	24	36.4	9	5	US-09-641-528-46197
11	24	36.4	9	5	US-09-641-528A-16119
12	24	36.4	9	5	US-09-641-528A-46197
13	24	36.4	10	5	US-09-641-528-23707
14	24	36.4	10	5	US-09-641-528A-23707
15	24	36.4	10	6	US-10-283-903-34
16	24	36.4	11	5	US-09-641-528-16120
17	24	36.4	11	5	US-09-641-528-19850
18	24	36.4	11	5	US-09-641-528A-16120
19	24	36.4	11	5	US-09-641-528A-19850
20	24	36.4	13	1	PCT-US02-16164-74
21	24	36.4	13	1	PCT-US02-16164-137
22	24	36.4	13	1	PCT-US02-16164-187
23	24	36.4	13	1	PCT-US02-16164-237
24	23	34.8	9	1	PCT-US02-31642-330
25	23	34.8	9	6	US-10-283-903-17
26	23	34.8	9	6	US-10-283-903-146

ALIGNMENTS

RESULT 1

US-09-641-528-13672
; Sequence 13672, Application US/09641528
; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528

; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705

; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13672

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus

US-09-641-528-13672

Query Match

Best Local Similarity

Matches

39.4%; Score 26; DB 5; Length 9;

6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SPXFTKFDL 13

DB 1 SFNDTTFDL 9

RESULT 2

US-09-641-528-45663

; Sequence 45663, Application US/09641528

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS

; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

; FILE REFERENCE: 2060.0100001

Sequence 163, App
Sequence 233, App
Sequence 330, App
Sequence 87, Appl
Sequence 100, App
Sequence 387, App
Sequence 674, App
Sequence 435, App
Sequence 3167, App
Sequence 3167, App
Sequence 674, App
Sequence 21628, A
Sequence 30698, A
Sequence 34208, A
Sequence 41604, A
Sequence 674, App
Sequence 21628, A
Sequence 30698, A
Sequence 34208, A

27 23 34.8 9 6 US-10-283-903-163
28 23 34.8 9 6 US-10-283-903-233
29 23 34.8 9 6 US-10-264-309-330
30 23 34.8 10 6 US-10-283-903-87
31 23 34.8 10 6 US-10-283-903-100
32 23 34.8 10 6 US-10-283-903-387
33 23 34.8 10 6 US-10-283-903-674
34 23 34.8 12 5 US-09-676-475A-435
35 22 33.3 5 1 PCT-US02-36496-3167
36 22 33.3 5 6 US-10-293-418-3167
37 22 33.3 8 5 US-09-641-528-674
38 22 33.3 8 5 US-09-641-528-21628
39 22 33.3 8 5 US-09-641-528-30698
40 22 33.3 8 5 US-09-641-528-34208
41 22 33.3 8 5 US-09-641-528-41604
42 22 33.3 8 5 US-09-641-528A-674
43 22 33.3 8 5 US-09-641-528A-21628
44 22 33.3 8 5 US-09-641-528A-30698
45 22 33.3 8 5 US-09-641-528A-34208

; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45663
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-45663

Query Match 39.4%; Score 26; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
|||
Db 1 SFNDTTFDL 9

RESULT 3

US-09-641-528A-13672
; Sequence 13672, Application US/09641528A

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A

; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705

; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13672

; LENGTH: 9

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus

US-09-641-528A-13672

Query Match 39.4%; Score 26; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
|||
Db 1 SFNDTTFDL 9

RESULT 4

US-09-641-528A-45663

; Sequence 45663, Application US/09641528A

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A

; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45663
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-45663

Query Match 39.4%; Score 26; DB 5; Length 9;
Best Local Similarity 66.7%; Pred. No. 4e+05;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
|||
Db 1 SFNDTTFDL 9

RESULT 5

US-09-641-528-21738

; Sequence 21738, Application US/09641528

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528

; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705

; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21738

; LENGTH: 10

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Peptide Derived from Human Papillomavirus

US-09-641-528-21738

Query Match 39.4%; Score 26; DB 5; Length 10;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
|||
Db 2 SFNDTTFDL 10

RESULT 6

US-09-641-528A-21738

; Sequence 21738, Application US/09641528A

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard

; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

; FILE REFERENCE: 2060.0100001

; CURRENT APPLICATION NUMBER: US/09/641,528A

; CURRENT FILING DATE: 2000-08-15

; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21738
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-21738

Query Match 39.4%; Score 26; DB 5; Length 10;
Best Local Similarity 66.7%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||| |||
Db 2 SFNDTTFDL 10

RESULT 7

US-09-641-528-18763
; Sequence 18763, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18763
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-18763

Query Match 39.4%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||| |||
Db 3 SFNDTTFDL 11

RESULT 8

US-09-641-528A-18763
; Sequence 18763, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705

; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18763
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-18763

Query Match 39.4%; Score 26; DB 5; Length 11;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||| |||
Db 3 SFNDTTFDL 11

RESULT 9

US-09-641-528-16119
; Sequence 16119, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16119
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-16119

Query Match 36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||| : |||
Db 1 SFTYPKYSL 9

RESULT 10

US-09-641-528-46197
; Sequence 46197, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; TITLE OF INVENTION: USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10

; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46197
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-46197

Query Match 36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||:|:|
Db 1 SFTYPKYSL 9

RESULT 11
US-09-641-528A-16119
; Sequence 16119, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16119
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-16119

Query Match 36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||:|:|
Db 1 SFTYPKYSL 9

RESULT 12
US-09-641-528A-46197
; Sequence 46197, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 46197
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-46197

Query Match 36.4%; Score 24; DB 5; Length 9;
Best Local Similarity 44.4%; Pred. No. 4e+05;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||:|:|
Db 1 SFTYPKYSL 9

RESULT 13
US-09-641-528-23707
; Sequence 23707, Application US/09641528
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 23707
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528-23707

Query Match 36.4%; Score 24; DB 5; Length 10;
Best Local Similarity 44.4%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 SFXFTKFDL 13
||:|:|
Db 2 SFTYPKYSL 10

RESULT 14
US-09-641-528A-23707
; Sequence 23707, Application US/09641528A
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard
; TITLE OF INVENTION: INDUCING CELLULAR IMMUNE RESPONSES TO HUMAN PAPILLOMAVIRUS
; FILE REFERENCE: 2060.0100001
; CURRENT APPLICATION NUMBER: US/09/641,528A
; CURRENT FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 51504
; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 23707
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide Derived from Human Papillomavirus
US-09-641-528A-23707

Query Match 36.4%; Score 24; DB 5; Length 10;
Best Local Similarity 44.4%; Pred. No. 72;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

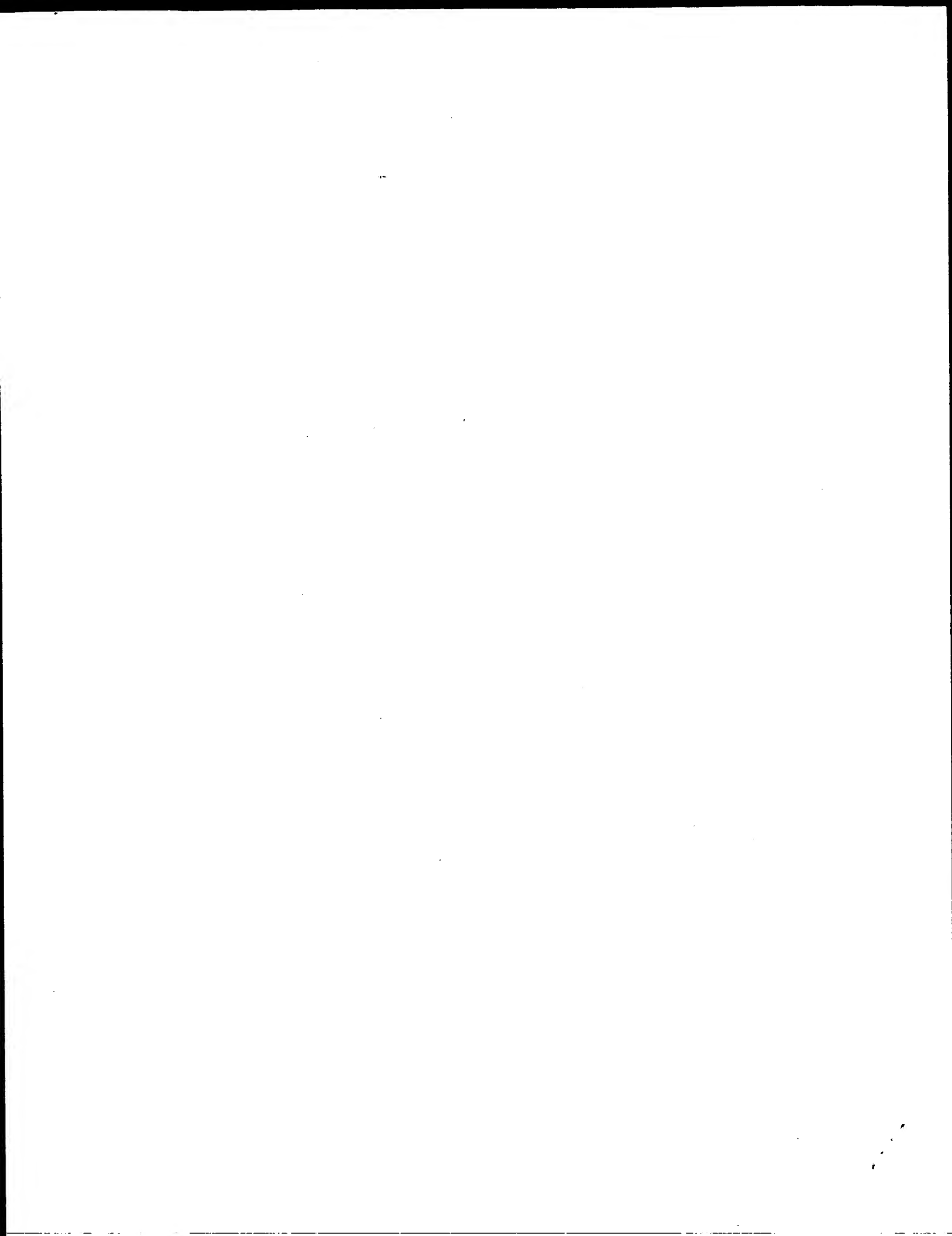
QY 5 SFETKFDL 13
||:|:|
Db 2 SFTYPKYSL 10

RESULT 15
US-10-305-231-34
; Sequence 34, Application US/10305231
; GENERAL INFORMATION:
; APPLICANT: Huse, William D.
; APPLICANT: Wu, Herren
; TITLE OF INVENTION: Anti-Alphav Beta3 Recombinant Human Antibodies, Nucleic
; TITLE OF INVENTION: Acids Encoding Same and Methods of Use
; FILE REFERENCE: P-IX 3536
; CURRENT APPLICATION NUMBER: US/10/305,231
; CURRENT FILING DATE: 2002-11-25
; PRIOR APPLICATION NUMBER: APPLICATION NUMBER: US/09/339,922
; PRIOR FILING DATE: CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-305-231-34

Query Match 36.4%; Score 24; DB 6; Length 10;
Best Local Similarity 37.5%; Pred. No. 72;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 6 FXETKFDL 13
| |:|:
Db 2 FTFSSYDM 9

Search completed: February 26, 2003, 15:01:16
Job time : 23 secs



GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 26, 2003, 14:55:45 ; Search time 14 Seconds
(without alignments)
29.423 Million cell updates/sec

Title: US-09-476-485A-31

Perfect score: 66

Sequence: 1 AQSLSFXFTKFDLD 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 103252

Minimum DB seq length: 0

Maximum DB seq length: 14

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA.*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	98.5	14	3	US-08-825-369A-2
2	55	83.3	12	3	US-08-825-369A-1
3	25	37.9	8	4	US-09-187-859-2462
4	25	37.9	8	4	US-09-187-859-2582
5	25	37.9	9	4	US-09-187-859-2463
6	25	37.9	9	4	US-09-187-859-2583
7	25	37.9	9	4	US-09-644-600-24
8	25	37.9	9	4	US-09-644-600-70
9	25	37.9	10	4	US-09-187-859-2464
10	25	37.9	10	4	US-09-187-859-2584
11	25	37.9	11	2	US-08-486-839-9
12	25	37.9	11	3	US-09-151-011-9
13	25	37.9	11	4	US-09-343-623-9
14	25	37.9	12	3	US-08-599-226-31
15	25	37.9	12	4	US-09-125-098-31
16	24	36.4	7	1	US-08-092-110A-4
17	24	36.4	7	1	US-08-273-474-4
18	24	36.4	7	4	US-08-935-100-4
19	24	36.4	7	5	PCT-US94-07881-4
20	24	36.4	8	2	US-08-350-260A-381
21	24	36.4	10	2	US-08-480-190-149
22	24	36.4	10	2	US-08-488-379-149
23	24	36.4	10	1	PCT-US93-07545-149
24	24	36.4	12	5	US-08-057-184-1
25	24	36.4	12	2	US-08-480-190-100
26	24	36.4	12	2	US-08-480-190-111
27	24	36.4	12	2	US-08-480-190-187

28	24	36.4	12	2	US-08-488-379-100	Sequence 100, App
29	24	36.4	12	2	US-08-488-379-111	Sequence 111, App
30	24	36.4	12	2	US-08-488-379-187	Sequence 187, App
31	24	36.4	12	5	PCT-US93-07545-100	Sequence 100, App
32	24	36.4	12	5	PCT-US93-07545-111	Sequence 111, App
33	24	36.4	12	5	PCT-US93-07545-187	Sequence 187, App
34	24	36.4	13	1	US-08-057-184-2	Sequence 2, Appl
35	24	36.4	13	4	US-09-171-878-19	Sequence 19, Appl
36	24	36.4	13	4	US-09-194-285-56	Sequence 56, Appl
37	24	36.4	14	2	US-08-480-190-110	Sequence 110, App
38	24	36.4	14	2	US-08-488-379-110	Sequence 110, App
39	24	36.4	14	5	PCT-US93-07545-110	Sequence 110, App
40	23	34.8	7	4	US-09-147-933-14	Sequence 14, Appl
41	23	34.8	8	5	PCT-US95-03236-19	Sequence 19, Appl
42	23	34.8	12	1	US-08-704-170-54	Sequence 54, Appl
43	23	34.8	12	4	US-09-258-754-435	Sequence 435, App
44	23	34.8	12	4	US-09-042-107-435	Sequence 435, App
45	23	34.8	12	5	PCT-US94-02631-54	Sequence 54, Appl

ALIGNMENTS

RESULT 1
US-08-825-369A-2
; Sequence 2, Application US/08825369A
; Patent No. 6084060
; GENERAL INFORMATION:
; APPLICANT: Moore
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
; TITLE OF INVENTION: PROGENITOR CELLS
; NUMBER OF SEQUENCE: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: USA
; ZIP: 11753
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/825,369A
; FILING DATE: March 28, 1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Dea, Sean W.
; REGISTRATION NUMBER: 37690
; REFERENCE/DOCKET NUMBER: 381-21 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-825-369A-2

Query Match 98.5%; Score 65; DB 3; Length 14;
Best Local Similarity 92.9%; Pred. No. 1.9e-06;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXFTKFDLD 14

Db 1 AQSLSFXFTKFDLD 14

RESULT 2

US-08-825-369A-1
 ; Sequence 1, Application US/08825369A
 ; Patent No. 6084060
 ; GENERAL INFORMATION:
 ; APPLICANT: Moore
 ; TITLE OF INVENTION: COMPOSITION AND METHOD FOR PRESERVING
 ; TITLE OF INVENTION: PROGENITOR CELLS
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hoffmann & Baron, LLP
 ; STREET: 350 Jericho Turnpike
 ; CITY: Jericho
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 11753
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
 ; COMPUTER: IBM compatible
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WordPerfect
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/825,369A
 ; FILING DATE: March 28, 1997
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: O'Dea, Sean W.
 ; REGISTRATION NUMBER: 37690
 ; REFERENCE/DOCKET NUMBER: 381-21 CIP
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (516) 822-3550
 ; TELEFAX: (516) 822-3582
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 12 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-825-369A-1

Query Match 83.3%; Score 55; DB 3; Length 12;
 Best Local Similarity 91.7%; Pred. No. 0.00013;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 AQSLSFXTKFD 12
 ||||| |||||
 DB 1 AQSLSFSTKFD 12

RESULT 3
 US-09-187-859-2462
 ; Sequence 2462, Application US/09187859A
 ; Patent No. 6358920
 ; GENERAL INFORMATION:
 ; APPLICANT: Blaschuk, Orest W.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
 ; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
 ; FILE REFERENCE: 100086.407C1
 ; CURRENT APPLICATION NUMBER: US/09/187,859A
 ; CURRENT FILING DATE: 1998-11-06
 ; NUMBER OF SEQ ID NOS: 4052
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2462
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Representative cyclic modulating agent based on
 ; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
 ; US-09-187-859-2462

Query Match 37.9%; Score 25; DB 4; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.9e+05;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 10 KFDLD 14
 |||:|
 DB 1 KFDID 5
 RESULT 4
 US-09-187-859-2582
 ; Sequence 2582, Application US/09187859A
 ; Patent No. 6358920
 ; GENERAL INFORMATION:
 ; APPLICANT: Blaschuk, Orest W.
 ; APPLICANT: Gour, Barbara J.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
 ; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
 ; FILE REFERENCE: 100086.407C1
 ; CURRENT APPLICATION NUMBER: US/09/187,859A
 ; CURRENT FILING DATE: 1998-11-06
 ; NUMBER OF SEQ ID NOS: 4052
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2582
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Representative cyclic modulating agent based on
 ; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
 ; US-09-187-859-2582

Query Match 37.9%; Score 25; DB 4; Length 8;
 Best Local Similarity 80.0%; Pred. No. 1.9e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
 |||:|
 DB 1 KFDID 5

RESULT 5
 US-09-187-859-2463
 ; Sequence 2463, Application US/09187859A
 ; Patent No. 6358920
 ; GENERAL INFORMATION:
 ; APPLICANT: Blaschuk, Orest W.
 ; APPLICANT: Gour, Barbara J.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
 ; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
 ; FILE REFERENCE: 100086.407C1
 ; CURRENT APPLICATION NUMBER: US/09/187,859A
 ; CURRENT FILING DATE: 1998-11-06
 ; NUMBER OF SEQ ID NOS: 4052
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2463
 ; LENGTH: 9
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Representative cyclic modulating agent based on
 ; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
 ; US-09-187-859-2463

Query Match 37.9%; Score 25; DB 4; Length 9;
 Best Local Similarity 80.0%; Pred. No. 1.9e+05;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 10 KFDLD 14
 |||:|
 DB 1 KFDID 5

RESULT 6
 US-09-187-859-2583

```

; Sequence 2583, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2583
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-187-859-2583

Query Match          37.9%; Score 25; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14
| | | |
Db 1 KFDID 5

RESULT 7
US-09-644-600-24
; Sequence 24, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 24
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 257-265 of the TADG-15 protein
US-09-644-600-24

Query Match          37.9%; Score 25; DB 4; Length 9;
Best Local Similarity 80.0%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SFXTKFDL 13
| | | |
Db 1 SLTFRSFDL 9

RESULT 8
US-09-644-600-70
; Sequence 70, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotochi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas

```

```

; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 70
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 257-265 of the TADG-15 protein
US-09-644-600-70

Query Match          37.9%; Score 25; DB 4; Length 9;
Best Local Similarity 55.6%; Pred. No. 1.9e+05;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SFXTKFDL 13
| | | |
Db 1 SLTFRSFDL 9

RESULT 9
US-09-187-859-2464
; Sequence 2464, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2464
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-187-859-2464

Query Match          37.9%; Score 25; DB 4; Length 10;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14
| | | |
Db 1 KFDID 5

RESULT 10
US-09-187-859-2584
; Sequence 2584, Application US/09187859A
; Patent No. 6358920
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING NONCLASSICAL
; TITLE OF INVENTION: CADHERIN-MEDIATED FUNCTIONS
; FILE REFERENCE: 100086.407C1
; CURRENT APPLICATION NUMBER: US/09/187,859A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 4052
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2584
; LENGTH: 10

```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Representative cyclic modulating agent based on
; OTHER INFORMATION: PB-cadherin cell adhesion recognition sequence
US-09-187-859-2584

Query Match      37.9%; Score 25; DB 4; Length 10;
Best Local Similarity 80.0%; Pred. No. 57;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 10 KFDLD 14
   |||:|
Db 1 KFDID 5

RESULT 11
US-08-486-839-9
; Sequence 9, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-486-839-9

Query Match      37.9%; Score 25; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14
   | ||||
Db 1 FDGFDLD 7

RESULT 12
US-09-151-011-9
; Sequence 9, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:

```

```

; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-151-011-9

Query Match      37.9%; Score 25; DB 3; Length 11;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 FTKFDLD 14
   | ||||
Db 1 FDGFDLD 7

RESULT 13
US-09-343-623-9
; Sequence 9, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO

US-09-343-623-9

Query Match 37.9%; Score 25; DB 4; Length 11;
Best Local Similarity 71.4%; Pred No. 63;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 FTKFDLD 14
| | | | |
Db 1 FDGFDLD 7

RESULT 14

US-08-599-226-31
Sequence 31, Application US/08599226
Patent No. 6090382

GENERAL INFORMATION:

APPLICANT: Salfeld, Jochen G.
APPLICANT: Allen, Deborah J.
APPLICANT: Hoogenboom, Hendricus R.J.M.
APPLICANT: Kaymakcalan, Zehra
APPLICANT: Labkovsky, Boris
APPLICANT: Mankovich, John A.
APPLICANT: McGuinness, Brian T.
APPLICANT: Roberts, Andrew J.
APPLICANT: Sakorafas, Paul
APPLICANT: Schoenhaut, David
APPLICANT: Vaughan, Tristan J.
APPLICANT: White, Michael
APPLICANT: Wilton, Andrew J.

TITLE OF INVENTION: Human Antibodies that Bind Human TNFa

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/599,226

FILING DATE: 08-FEB-1996

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: DeConti, Giulio A., Jr.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: BBI-043

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
US-08-599-226-31

Query Match 37.9%; Score 25; DB 3; Length 12;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 SFXTKFDLD 14
| | | | |
Db 2 SYLSTFSLD 11

RESULT 15

US-09-125-098-31

Sequence 31, Application US/09125098

Patent No. 6258562

GENERAL INFORMATION:

APPLICANT: Salfeld, Jochen G.
APPLICANT: Allen, Deborah J.
APPLICANT: Hoogenboom, Hendricus R.J.M.
APPLICANT: Kaymakcalan, Zehra
APPLICANT: Labkovsky, Boris
APPLICANT: Mankovich, John A.
APPLICANT: McGuinness, Brian T.
APPLICANT: Roberts, Andrew J.
APPLICANT: Sakorafas, Paul
APPLICANT: Schoenhaut, David
APPLICANT: Vaughan, Tristan J.
APPLICANT: White, Michael
APPLICANT: Wilton, Andrew J.

TITLE OF INVENTION: Human Antibodies that Bind Human TNFa

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/125,098

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/599,226

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: DeConti, Giulio A., Jr.

REGISTRATION NUMBER: 31,503

REFERENCE/DOCKET NUMBER: BBI-043

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-09-125-098-31

Query Match

37.9%; Score 25; DB 4; Length 12;

Best Local Similarity 50.0%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches

4; Indels 0; Gaps 0;

Qy 5 SFXTKFDLD 14

l: | | | |

Db 2 SYLSTSFSLD 11

Search completed: February 26, 2003, 14:58:27
Job time : 14 secs